

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHQESTXGRSRGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	93.5	2248	2	A35938
2	84	91.3	416	2	A32947
3	75	81.5	591	2	A45135
4	58	63.0	506	1	W2WT47
5	55	59.8	822	2	T51049
6	49	53.3	849	2	A96592
7	48	52.2	136	2	T35632
8	47	51.1	471	2	T33997
9	47	51.1	825	2	JC4163
10	47	51.1	836	2	G84727
11	46	50.0	306	2	T21320
12	46	50.0	312	2	A31846
13	46	50.0	399	2	T46259
14	45.5	49.5	135	2	A46398
15	44	47.8	843	1	JC7219
16	44	47.8	829	2	JDVLVR
17	44	47.8	1804	2	T34518
18	44	47.8	2420	2	A84652
19	43.5	47.3	463	2	T51194
20	43	46.7	123	2	T16234
21	43	46.7	141	2	T39554
22	43	46.7	531	2	D86242
23	43	46.7	745	2	G01025
24	43	46.7	774	2	I48609
25	43	46.7	800	2	T02852
26	43	46.7	827	2	T13468
27	43	46.7	842	1	JDVLVS
28	43	46.7	1119	2	T16720
29	43	46.7	1829	2	T35681

30	42.5	46.2	286	2	S50855	neurotrophin-6 - s
31	42	45.7	218	2	S58769	brain-specific pro
32	42	45.7	638	2	I53169	cytokereatin 2 - hu
33	42	45.7	694	2	S71786	wingless receptor
34	42	45.7	778	2	T44761	probable preprotet
35	42	45.7	808	2	F70720	probable transloca
36	42	45.7	1015	2	T42013	frequency clock pr
37	42	45.7	1507	2	B47328	natural killer cel
38	41.5	45.1	553	2	T27245	hypothetical prote
39	41.5	45.1	568	2	H88904	protein Y57G11C.9
40	41.5	45.1	659	2	T27246	hypothetical prote
41	41	44.6	176	2	F71370	probable single-st
42	41	44.6	289	2	AD3394	lipoprotein releas
43	41	44.6	315	2	S53589	S012 protein - Yea
44	41	44.6	373	2	T02976	probable DNA bindi
45	41	44.6	482	2	C86322	hypothetical prote

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GNN>
A:Cross-references: GB:U02929
A:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-568/Region: profilaggrin repeat
F:570-893/Region: profilaggrin repeat
F:1074-1397/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 93.5%; Score 86; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 7.3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTXGRSRGRSGSGS 19
DB 551 SHQESTXGRSRGRSGSGS 569
|||||:|||||
|||:|||||

RESULT 2
A32947
filaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A:Title: Characterization of a cDNA clone encoding human filaggrin and localization of tr
A:Reference number: A32947; MUID:89296901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M4355; NID:g182604; PIND:AAA52454.1; P1D:g182605
A>Note: The authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:

CjAccession: T35632
R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T35632
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-136 <SEB>
A:Cross-references: EMBL:AL079356; PIDN:CAB45627.1; GSPDB:GN00070; SCOPDB:SC6G9.36C
C:Genetics:
A:Gene: SCOPDB:SC6G9.36C
C:Superfamily: Streptomyces coelicolor probable transposase SC6G9.36C

Query Match 52.2%; Score 48; DB 2; Length 136;
Best Local Similarity 47.1%; Pred. No. 2.8;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 17
|||:|||||:
Db 111 AHQHAAGAKKRGRTGA 127

RESULT 8
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
CjSpecies: Caenorhabditis elegans
CjDate: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
CjAccession: T33997
R:Pauley, A.; Scheel, P.; Harper, M.
Submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: EMBL:AF125964; PIDN:AD14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol NZ; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
Best Local Similarity 52.6%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 19
|:|:|||||
Db 167 SSRSPGRGRGGRSGS 185

RESULT 9
JC4163
DNA-binding protein 5E5 - rat
CjSpecies: Rattus norvegicus (Norway rat)
CjDate: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
CjAccession: JC4163; PC4040
R:Suzuki, E.; Kojima, N.; Yoshimura, K.; Uyemura, K.; Odata, K.; Akagawa, K.
J. Biochem. 118, 122-128, 1995
A:Title: Cloning and sequence analysis of cDNA for a possible DNA-binding protein 5E5 in
A:Reference number: JC4163; MUID:96015159; PMID:8537300
A:Accession: JC4163
A:Molecule type: mRNA
A:Residues: 1-825 <SUZ>
A:Cross-references: DDBJ:D37934; NID:G531260; PIDN:BA07153.1; PID:G531261
A:Experimental source: brain
A:Accession: PC4040
A:Molecule type: protein
A:Residues: 230-455 <SUZ>
CjComment: This protein has an abundance of arginine, a glycine-rich region and a prolin
CjSuperfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
CjKeywords: nerve; phosphoprotein

F:436-443/Region: nuclear location signal
F:722-731/Region: proline cluster
F:62,258,345,360,404,413,570,613,635,752,820/Binding site: phosphate (Ser) (covalent) (b)

Query Match 51.1%; Score 47; DB 2; Length 825;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGRSGS 19
|||:|||||
Db 589 GRGRGRSGRSGS 600

RESULT 10
G84727
probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
CjSpecies: Arabidopsis thaliana (mouse-ear cress)
CjDate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
CjAccession: G84727
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Dmayem, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84727
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-836 <STO>
A:Cross-references: GB:AE02093; NID:94263718; PIDN:AAD15404.1; GSPDB:GN00139
C:Genetics:
A:Gene: AT2G32000
A:Map position: 2

Query Match 51.1%; Score 47; DB 2; Length 836;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGRSGS 19
|||:|||||
Db 813 GRGRGRSGRSGS 824

RESULT 11
T21220
hypothetical protein F21H7.5 - Caenorhabditis elegans
CjSpecies: Caenorhabditis elegans
CjDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
CjAccession: T21220
R:Gardner, A.
Submitted to the EMBL Data Library, March 1997
A:Reference number: Z19393
A:Accession: T21220
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-306 <WIL>
A:Cross-references: EMBL:T29379; PIDN:CAB07588.1; GSPDB:GN00023; CESP:F21H7.5
A:Experimental source: clone F21H7
C:Genetics:
A:Gene: CESP:F21H7.5
A:Map position: 5
A:Introns: 28/3; 146/3; 232/3
CjSuperfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 50.0%; Score 46; DB 2; Length 306;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGS 18
|||:|||||
Db 48 SHRSSKTKKSGKSGS 65

F;12-74/Domain: ribonucleoprotein repeat homology <RMR3>

Search completed: September 28, 2004, 06:14:56
Job time : 16.0625 secs

F;12-74/Domain: ribonucleoprotein repeat homology <RMR3>


```

DR PFAM, PF03516; Filaggrin; 2.
DR PRINTS; PR00487; FILAGRIN.
KM Phosphorylation; Cytullination; Developmental protein.
FT NON TER 1
SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
Best Local Similarity 89.5%; Pred. No. 4.9e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SH0ESTXGRSGRSGSGS 19
    |||||
    7 SH0ESTXGRSGRSGSGS 25

RESULT 2
ID VE2_HPV47 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
OS Human Papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9028161; Pubmed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -.
DR PIR; D35324; W2WL47.
DR HSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C_1.
DR Pfam; PF00508; E2_N_1.
DR ProDom; PD006678; E2_C_1.
DR ProDom; PD006678; E2_N_1.
DR Early protein; Transcription regulation; Activator; DNA-binding;
KM Trans-acting factor; DNA replication; Repressor protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF8725065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.074;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SH0ESTXGRSGRSGSGS 19

```

```

DB 342 SREGNTRGRGRGRGRGS 360

RESULT 3
ID SES_RAT STANDARD; PRT; 825 AA.
AC O63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SES antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistax; TISSUE=Brain;
RX MEDLINE=9601519; Pubmed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
protein SES in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D37934; BA07153.1; -.
DR PIR; U04163; UC4163.
KM DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD55BDF CRC64;

Query Match 51.1%; Score 47; DB 1; Length 825;
Best Local Similarity 75.0%; Pred. No. 6.7;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 GRSRGRSGRSGS 19
    |||||
    DB 589 GRGRGRGRGRGS 600

RESULT 4
ID RBPI_DROME STANDARD; PRT; 135 AA.
AC Q02427; Q26271; Q9VGM8;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein 1.
GN RBPI OR CG17136.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=94040720; Pubmed=1340470;
RA Kim Y.-J., Zuo P., Manly J.L., Baker B.S.;
RT "The Drosophila RNA-binding protein RBPI is localized to
transcriptionally active sites of chromosomes and shows a functional
similarity to human splicing factor ASF/SF2.";
RL Genes Dev. 6:2565-2579(1992).
RN [2]

```

RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gockyne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Besu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclev J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spralling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN (3)
 RP SEQUENCE OF 14-52 FROM N.A.
 RX MEDLINE=93109300; PubMed=8417324;
 RA Kim Y.J., Baker B.S.;
 RT "Isolation of RRM-type RNA-binding protein genes and the analysis of
 their relatedness by using a numerical approach.";
 RL Mol. Cell. Biol. 13:174-183(1993).
 RN (4)
 RP FUNCTION;
 RX MEDLINE=95393975; PubMed=7664738;
 RA Heinrichs V., Baker B.S.;
 RT "The *Drosophila* SR protein RBP1 contributes to the regulation of
 doublesex alternative splicing by recognizing RBP1 RNA target
 sequences.";
 RL EMBO J. 14:3987-4000(1995).
 CC -1- FUNCTION: Contributes to the activation of female-specific DSX
 splicing in vivo by recognizing the RBP1 target sequences within
 the putative-rich polypyrimidine tract of the female-specific 3'
 splice site.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RBP1-A;
 CC IsoId=002427-1; Sequence=Displayed;
 CC Name=2; Synonyms=RBP1-B;
 CC IsoId=002427-2; Sequence=VSP_005817;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
 CC -1- PTM: Extensively phosphorylated on serine residues in the RS
 domain (Probable).
 CC -1- SIMILARITY: Belongs to the splicing factor SR family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; LC04929; AAA28850.1; -;
 DR EMBL; AE003668; AAF54555.1; -;
 DR EMBL; S51691; AAB24622.1; -;
 DR PIR; A46398; A46398.
 DR PIR; A48110; A48110.
 DR FLYBase; FBgn010252; Rbp1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 DR Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
 DR Phosphorylation.
 DR DOMAIN 11 84 RNA-BINDING (RRM).
 DR ARG/SER-RICH (RS DOMAIN).
 DR VAREPLIC 107 135 Missing (in isoform 2).
 DR /FTid=VSP_005817.
 FT CONFLICT 14 14 Y -> F (IN REF. 3).
 FT CONFLICT 65 65 R -> A (IN REF. 1).
 SQ SEQUENCE 135 AA; 15446 MW; 896DCE902518D991 CRC64;
 Query Match 49.5%; Score 45.5; DB 1; Length 135;
 Best Local Similarity 52.2%; Pred. No. 1.7;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
 QY 4 ESTXGRSR-----GRSGRSGS 19
 Db 80 EMSGRSHRRRGEGSGSGRGS 102
 RESULT 5
 DRPOL_HPBVR STANDARD; PRT; 843 AA.
 AC P03157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7), RNA-
 DE directed DNA polymerase (EC 2.7.7.49), Ribonuclease H (EC 3.1.26.4)].
 GN Hepatitis B virus (subtype adr).
 OS Hepatitis B virus (subtype adr).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=106820;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8316919; PubMed=6300776;
 RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioaka K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus
 RT DNA; subtype adr and adr.";
 RL Nucleic Acids Res. 11:1747-1757(1983).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonocester.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00867; -; NOT_ANNOTATED_CDS.
 DR PIR; A00704; JDIIVLR.

```

DR InterPro: IPR001462; DNAPol_viral_C.
DR InterPro: IPR000201; DNAPol_viral_N.
DR InterPro: IPR000477; RVase.
DR Pfam: PF00336; DNA_pol_viral_C; 1.
DR Pfam: PF00242; DNA_pol_viral_N; 1.
DR Pfam: PF00078; rvt1_2.
DR ProDom: PD000814; DNAPol_viral_C; 1.
DR TransErase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KM HydroLase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C48B CRC64;

OY Query Match 47.8%; Score 44; DB 1; Length 843;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

DQ 3 QESTXGRSGRSGSGS 19
DQ 226 QPQGSVARKSGRSGS 242

RESULT 6
LSM4_CAEEL STANDARD; PRT; 123 AA.
AC Q19952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.
CS Caeothabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabdioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Pauley A.; (MAR-1995) to the EMBL/GenBank/DBJ databases
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -1- SUBUNIT: LSM subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20864; AAC46661.1; -
DR PIR: T16234; T16234.
DR WormPep: F32A5.7; CE01277.
DR InterPro: IPR006649; snRNP.
DR InterPro: IPR001163; snRNP_Sm.
DR Pfam: PF01423; LSM; 1.
DR ProDom: PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
DR Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
KW RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

OY Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 3.8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

DQ 3 QESTXGRSGRSGSGS 18
DQ 93 RQSGRGRGRGRGRG 108

```

```

RESULT 7
PEXD_PICPA STANDARD; PRT; 380 AA.
AC Q92266;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Peroxisomal membrane protein PEX13 (Peroxin-13).
CN PEX13.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4922;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=97011155; PubMed=8958165;
RA Gould S.J., Kalish U.E., Morrell J.C., Bjorkman J., Urquhart A.J.,
RA Crane D.I.;
RT "Pex13p is an SH3 protein of the peroxisome membrane and a docking
RT factor for the predominantly cytoplasmic Pts1 receptor."
RT J. Cell Biol. 135:85-95(1996).
CC -1- FUNCTION: Component of the peroxisomal translocation machinery
CC with PEX14 and PEX17. Interacts with the Pts1 receptor
CC (Pals10/PEX5). Involved in the import of Pts1 and Pts2 proteins.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative Initiation;
CC Comment=2 isoforms, PEX13L (shown here) and PEX13S, are produced
CC by alternative initiation;
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U70067; AAB09087.1; -
DR HSP: Q64010; ICKA.
DR InterPro: IPR007223; Peroxin-13_N.
DR InterPro: IPR001452; SH3.
DR Pfam: PF04088; Peroxin-13_N; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
DR Peroxisome; Transport; Protein transport; Transmembrane; SH3 domain;
KW Alternative initiation.
SQ SEQUENCE 380 AA; 40695 MW; EEBAB39F93BA832 CRC64;

OY Query Match 46.7%; Score 43; DB 1; Length 380;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

DQ 1 SHQESTXGRSGRSGSGS 18

```

```

Db          315 SNOESTWKCRRSDGKVG 332
      ||||| : ||| :
RESULT 8
DPOL_HPBV4
ID DPOL_HPBV4 STANDARD; PRT; 730 AA.
AC p12933;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Hepatitis B virus (subtype adr4).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10409;
RP SEQUENCE FROM N.A.
RX MEDLINE=83246570; Pubmed=6306594;
RA Fujiyama A., Miyahara A., Nozaki C., Yoneyama T., Ohnomo N.,
RA Matsubara K.,
RT Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT adr."
RT Nucleic Acids Res. 11:4601-4610(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01587; CA25746.1; -
DR InterPro; IPR001462; DNADPOL_viral_C.
DR InterPro; IPR000201; DNADPOL_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C_1.
DR Pfam; PF00242; DNA_pol_viral_N_1.
DR ProDom; PD000814; DNADPOL_viral_C_1.
DR Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KM Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 730 AA; 82207 MW; 898A8BB5D0496F2 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 730;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSRGSGS 19
      |||||
Db          233 ARKSGRSGS 242

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93364122; Pubmed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk", a protein kinase with homologs in yeast maps to mouse
RT chromosome 19."
GN Mamm. Genome 4:401-403(1993).
RL -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70764; CA50040.1; -
DR PIR; I48609; I48609.
DR HSSP; Q63450; I406.
DR MGD; MGI:99638; Mark2.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot Kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot Kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 53 304
FT NP BIND 67
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7B5443493A CRC64;

Query Match 46.7%; Score 43; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGSRGSRGSGS 19
      |||||
Db          614 SPSGHSQGRGSGS 628
      |||||
RESULT 10
DPOL_HPBVM
ID DPOL_HPBVM STANDARD; PRT; 842 AA.
AC P31870; Q67861; Q67866;
DT 01-JUN-1993 (Rel. 26, Created)
DT 01-JUN-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Hepatitis B virus (subtype adr / mutant).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31512;
RP SEQUENCE FROM N.A.
RX MEDLINE=89183619; Pubmed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
RL Nucleic Acids Res. 17:2124-2124(1989).

```

```

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5' -
CC phosphomonester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14193; CA32399.1; -
DR EMBL; X14193; CA32405.1; ALT_TERM.
DR PIR; S04568; JDLVLS.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RTase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; Irt; 2.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR TransErase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Rnuclease; DNA replication; DNA-binding.
SQ SEQUENCE 842 AA; 94545 MW; 137FE37246BA5CC6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 842;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242

RESULT 11
ID P25A_BOVIN STANDARD; PRT; 218 AA.
AC Q2957; O9TRW3; O9TRW4;
DT 01-NOV-1987 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 25 kDa brain-specific protein (p25-alpha).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95375016; PubMed=7647094;
RA Shitarauchi A., Sato S., Oomori A., Ishiguro K., Uchida T.,
RA Imahori K.;
RT "cDNA cloning of a novel brain-specific protein p25."
RL Biochim. Biophys. Acta 1251:66-68(1995).
[2]
RP SEQUENCE OF 1-27, AND PHOSPHORYLATION.
RC TISSUE=Brain;
RX MEDLINE=91372400; PubMed=1909972;
RA Takahashi M., Tomizawa K., Ishiguro K., Sato K., Oomori A., Sato S.,
RA Shitarauchi A., Uchida T., Imahori K.;
RT "A novel brain-specific 25 kDa protein (p25) is phosphorylated by a
RT Ser/Thr-Pro kinase (TPK II) from tau protein kinase fractions."
RL FEBS Lett. 289:37-43(1991).
CC -1- TISSUE SPECIFICITY: Brain specific.
CC -1- PTM: Phosphorylated by TPK II.
CC -1- SIMILARITY: Belongs to the p25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85738; CA59741.1; -
DR PIR; S58769; S58769.
DR InterPro; IPR008507; P25-alpha.
DR Pfam; PF05517; p25-alpha; 1.
KW Phosphorylation.
SQ SEQUENCE 218 AA; 23472 MW; 55F5DAB42DC3A638 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 218;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 9; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 1 SHOE-STXGRSGRSGR 16
DB 174 SHKEFPDPSGRGKGRAGR 191

RESULT 12
ID K220_HUMAN STANDARD; PRT; 638 AA.
AC Q01546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2p) (K2p)
DE (CK 2p).
GN KR12P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93114504; PubMed=1282112;
RA Collin C., Ohnayeoun J.P., Grund C., Franke W.W.;
RT "Suprabasal marker proteins distinguishing keratinizing squamous
RT epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium
RT and epidermis are different."
RL Differentiation 51:137-148(1992).
CC -1- FUNCTION: Probably contributes to terminal cornification.
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC keratinocytes.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC (neutral to basic; 56-70 kDa) [K1 to K8].
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M99063; AAA3746.1; -
DR PIR; I53169; I53169.
DR GO; GO:0005882; C:intermediate filament; NMS.
DR GO; GO:0005198; F:structural molecule activity; NMS.
DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; NMS.
DR InterPro; IPR001664; IF.
DR InterPro; IPR002957; Keratin_I.
DR InterPro; IPR003054; Keratin_II.
DR Pfam; PF00038; filament; 1.
DR PRINTS; PR01246; TYPE1KERATIN.
DR PRINTS; PR01276; TYPE2KERATIN.
DR PROSITE; PS00226; IF_1.
KW Intermediate filament; Coiled coil; Keratin; Phosphorylation.

```

```

FT DOMAIN 1 182 HEAD.
FT DOMAIN 183 492 ROD.
FT DOMAIN 493 638 TAIL.
FT DOMAIN 183 218 COIL 1A.
FT DOMAIN 219 237 LINKER 1.
FT DOMAIN 238 329 COIL 1B.
FT DOMAIN 330 353 LINKER 12.
FT DOMAIN 354 492 COIL 2.
FT MOD RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 638 AA; 65871 MW; 9B743AB872076AF CRC64;

Query Match 45.7%; Score 42; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGRS 17
Db 7 KKSFGSRSGFSGRS 21

RESULT 13
FR22 DROME STANDARD; PRT; 694 AA.
AC Q9VWX3; Q94916; Q9VWX2;
DT 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Fritzed protein 2 precursor (Fritzed-2) (dfz2).
GN FZ2 OR CG9739/CG14083.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
RX MEDLINE=96353971; PubMed=8717036;
RA Bharat P., Brink M., Samos C.H., Heish J.C., Wang Y., Macke J.P.,
RA Andrew D., Nathans J., Nuss R.;
RT "A new member of the fritzed family from Drosophila functions as a
RT WINGLESS receptor.";
RL Nature 382:225-230(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceolnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Achyavani A., An H.-J., Andrews-Flankoch C., Baldwin D.,
RA Beeson R.M., Basu A., Baxendale U., Bayraktiroglu U., Beasley E.M.,
RA Borrova D., Botchan M.R., Boucek J., Brokstein P., Broctlier P.,
RA Chertis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Burris J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.B., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemsinn J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz J., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mironov G., Mishina N.V., Modyaty C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

-1- FUNCTION: Receptor for Wnt proteins. Most of fritzed receptors
are coupled to the beta-catenin canonical signaling pathway, which
leads to the activation of dishevelled proteins, inhibition of
GSK-3 kinase, nuclear accumulation of beta-catenin and activation
of Wnt target genes. A second signaling pathway involving PKC and
calcium fluxes has been seen for some family members, but it is
not yet clear if it represents a distinct pathway or if it can be
integrated in the canonical pathway, as PKC seems to be required
for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
to involve interactions with G-proteins. Required to coordinate
the cytoskeletons of epidermal cells to produce a parallel array
of cuticular hairs and bristles.

-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).

-1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
between 15 and 70 per cent of egg length, including the
invaginating cells of the ventral furrow. Stripe pattern is
emerging by early stage 8. From stage 9 and continuing throughout
embryogenesis, expression is seen in the developing CNS. At stage
10, expressed in 15 stripes in the presumptive head and trunk
regions, in the posterior midgut primordium, in a subset of cells
of anterior midgut invagination and in the procephalic lobe. At
stage 12, expression declines in epidermis and increases in the
midgut and visceral mesoderm. At stage 17, only expressed in the
CNS, hindgut and dorsal vessel.

-1- DOMAIN: Lys-Thr-X-X-Tyr motif is involved in the activation of
the Wnt/beta-catenin signaling pathway (By similarity).

-1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.

-1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
family.

-1- SIMILARITY: Contains 1 fritzed (Fz) domain.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U65589; AAC47273.1; -
DR EMBL: AB003518; AAF9385.2; -
DR PIR: S71786; S71786.
DR FLYbase; FBgn0016797; Fz2.
DR GO: GO:0016021; C: integral to membrane; NAS.
DR GO: GO:0017147; F: Wnt-protein binding; IDA.
DR GO: GO:0006928; P: cell motility; IMP.
DR GO: GO:0007163; P: establishment and/or maintenance of cell po. . .; ISS.
DR GO: GO:0008585; P: female gonad development; IMP.
DR GO: GO:0016055; P: Wnt receptor signaling pathway; IDA.
DR InterPro; IPR000539; Fritzed.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR000832; GPCR secretin.
DR Pfam; PF01534; Fritzed; I.
DR Pfam; PF01392; Fz; 1.
DR PRINTS; PR00489; FRITZLED.
DR SMART; SM00063; FRI; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50261; G-PROTEIN RECEPTOR; Transmembrane;
DR Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
DR Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 694 FRITZLED PROTEIN 2.

```



```

RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsenpe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrett B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RL "The complete genome sequence of Mycobacterium bovis."
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- FUNCTION: Involved in protein export. Interacts with the secY/secE
CC subunits. SecA has a central role in coupling the hydrolysis of
CC ATP to the transfer of pre-secretory periplasmic and outer
CC membrane proteins across the membrane (By similarity).
CC -1- SUBUNIT: Part of the prokaryotic protein translocation apparatus
CC which comprise secA, secB, secD, secE, secF, secG and secY
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (By
CC similarity).
CC -1- SIMILARITY: Belongs to the secA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z78020; CAB01458.1; -.
DR EMBL: AE007045; AAK46142.1; -.
DR EMBL: BX248340; CAD94555.1; -.
DR PIR: F70720; F70720.
DR TIGR: MT1869; -.
DR TubercuList; Rv1821; -.
DR InterPro; IPR001650; Helicase_C.
DR DR InterPro; IPR000185; SecA.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF01043; SecA protein; 1.
DR PRINTS; PR00906; SEC_A.
DR PROSITE; PS01312; SEC_A; 1.
KW Protein transport; ATP-binding; Membrane; Translocation; Transport;
KW Complete proteome.
FT NP BIND 139 146 ATP (POTENTIAL).
SQ SEQUENCE 808 AA; 88951 MW; 6AF6707CC3951D2F CRC64;

```

```

Query Match 45.7%; Score 42; DB 1; Length 808;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 HOESTYGRSRGRSGRG 18
DB 564 HTERLDNQLRGRARGG 580

```

Search completed: September 28, 2004, 06:05:45
Job time : 10.224 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHQESTYGRSGRSGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rvlinus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	97.8	797	4	Q16824	Q16824 homo sapien
2	86	93.5	990	4	Q15206	Q15206 homo sapien
3	84	91.3	1218	4	Q05331	Q05331 homo sapien
4	84	91.3	798	4	Q9H4U3	Q9H4U3 homo sapien
5	84	91.3	1084	4	Q01212	Q01212 homo sapien
6	75	81.5	591	4	Q01720	Q01720 homo sapien
7	75	81.5	687	4	Q9H4U2	Q9H4U2 homo sapien
8	70	76.1	465	4	Q03838	Q03838 homo sapien
9	69	75.0	322	4	Q05370	Q05370 homo sapien
10	55	59.8	822	3	Q9P312	Q9P312 neurospora
11	50	54.3	1284	5	Q9V9Y3	Q9V9Y3 drosophila
12	49	53.3	820	12	Q68375	Q68375 mouse cytom
13	49	53.3	849	10	Q9C7M2	Q9C7M2 arabidopsis
14	49	53.3	849	10	Q949M6	Q949M6 arabidopsis
15	49	53.3	1142	13	Q7SZV2	Q7SZV2 xenopus lae
16	48	52.2	481	16	Q7UUP9	Q7UUP9 rhodospirillum rubrum

17	48	52.2	1229	3	Q86ZH1	Q86ZH1 neurospora
18	47	51.1	471	5	Q9AYO	Q9AYO caenorhabditis
19	47	51.1	836	10	Q9SKZ9	Q9SKZ9 arabidopsis
20	46	50.0	306	5	Q45386	Q45386 caenorhabditis
21	46	50.0	399	4	Q9NTX9	Q9NTX9 homo sapien
22	46	50.0	436	16	Q82C67	Q82C67 streptococcus
23	46	50.0	467	4	Q9BW18	Q9BW18 homo sapien
24	46	50.0	700	13	Q42378	Q42378 brachydanio
25	46	50.0	715	13	Q8CFU0	Q8CFU0 brachydanio
26	46	50.0	715	13	Q8CGG8	Q8CGG8 brachydanio
27	46	50.0	716	13	Q42107	Q42107 brachydanio
28	46	50.0	737	4	Q9BQ39	Q9BQ39 homo sapien
29	46	50.0	787	5	Q9GSD4	Q9GSD4 plasmodium
30	46	50.0	843	12	Q9QNS2	Q9QNS2 hepatitis b
31	45.5	49.5	144	5	Q8T9K6	Q8T9K6 drosophila
32	45	48.9	96	15	Q8JRU5	Q8JRU5 human immun
33	45	48.9	96	15	Q9KRE7	Q9KRE7 human immun
34	45	48.9	96	15	Q9KRF4	Q9KRF4 human immun
35	45	48.9	96	15	Q8UBV4	Q8UBV4 human immun
36	45	48.9	96	15	Q8ADV5	Q8ADV5 human immun
37	45	48.9	243	10	Q9AX45	Q9AX45 oryza sativa
38	45	48.9	247	10	Q8GV23	Q8GV23 chlamydomonas
39	45	48.9	426	4	Q8WV76	Q8WV76 homo sapien
40	45	48.9	476	12	Q80890	Q80890 herpesvirus
41	45	48.9	652	5	Q77168	Q77168 apis mellifera
42	44.5	48.4	319	10	Q94AJ0	Q94AJ0 arabidopsis
43	44	47.8	96	15	Q8ADX3	Q8ADX3 human immun
44	44	47.8	229	11	Q9JW93	Q9JW93 mus musculus
45	44	47.8	259	4	Q86YA2	Q86YA2 homo sapien

ALIGNMENTS

RESULT 1

ID	Q16824	PRELIMINARY;	PRT;	797 AA.
AC	Q16824			
DT	01-NOV-1996	(TRENBLREL. 01, Created)		
DT	01-NOV-1996	(TRENBLREL. 01, Last sequence update)		
DT	01-UN-2003	(TRENBLREL. 24, Last annotation update)		
DE	Profilaggrin (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91064347; PubMed=2248957;			
RA	Gan S.O., McBride O.W., Idler W.W., Markova N., Steinhert P.M.;			
RT	"Organization, structure, and polymorphisms of the human profilaggrin			
RT	gene [published erratum appears in Biochemistry 1991 Jun			
RT	11;30(23):5814]."			
RL	Biochemistry 29:9432-9440(1990).			
DR	EMBL; M60502; AAA63248.1; -			
DR	GO; GO:0005198; F:structural molecule activity; IEA.			
DR	InterPro; IPR003303; Filaggrin.			
DR	Pfam; PF03516; Filaggrin; 4.			
DR	PRINTS; PR00487; FILAGGRIN.			
FT	NON_TER			
FT	SEQUENCE 797 AA; 85176 MW; 60E6184763BDA6B CRC64;			

Query Match 97.8%; Score 90; DB 4; Length 797;
Best Local Similarity 94.7%; Pred. No. 5.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTYGRSGRSGRSGS 19
DB 427 SHQESTYGRSGRSGRSGS 445

RESULT 2

Q15206
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FIG.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene [published erratum appears in Biochemistry 1991 Jun
11;30(23):5814]."
RT 11,30(23):5814.
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene".
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1; -;
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
Query Match 93.5%; Score 86; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 3.2e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SHORSTXGRSGRSGSGS 19
Db 551 SHORSTXGRSGRSGSGS 569
RESULT 3
005331
ID Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein."
RL Mol. Cell. Biol. 13:613-625(1993).
CC -I- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -I- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
CC -I- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1; -;
DR PIR; A48118; A48118.
DR HSSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00467; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;
Query Match 93.5%; Score 86; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CY 1 SHORSTXGRSGRSGSGS 19
Db 773 SHORSTXGRSGRSGSGS 791
RESULT 4
09H4U3
ID 09H4U3 PRELIMINARY; PRT; 798 AA.
AC 09H4U3;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE DJIAN1.1.2 (Profilaggrin 3' end) (Fragment).
GN FIG.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA laird G.;
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RX EMBL; AL356504; CAC13171.1; -;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 798
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 5.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 SHORSTXGRSGRSGSGS 19
Db 428 SHORSTXGRSGRSGSGS 446
RESULT 5
001212
ID 001212 PRELIMINARY; PRT; 1084 AA.

```

AC 001212; 003840;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride C.W., Idler W.W., Markova N., Steinhart P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; F:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR NON TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 91.3%; Score 84; DB 4; Length 1084;
Best Local Similarity 89.5%; Pred. No. 7.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 64 SHOESTGRSGRSGSGS 82

RESULT 6
Q01720
AC 001720; PRELIMINARY; PRT; 591 AA.
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Ntunukisiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus.";
RL J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CABP; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KM Polymorphism. 1
FT PROPEP 1 293
FT CHAIN 294 467
FT PROPEP 468 474
FT CHAIN 475 >591
FT CA_BIND 19 32
FT CA_BIND 62 73
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.0011;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
Db 449 SHOESTGRSGRSGRSGS 467

RESULT 7
Q9H4U2
AC Q9H4U2; PRELIMINARY; PRT; 687 AA.
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE DJL4N1.1.1 (profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN 1;
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 687;
Best Local Similarity 84.2%; Pred. No. 0.0012;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 SHOESTXGRSGRSGSGS 19
      ||||| |||||
DB      449 SHOESTXGRSGRSGSGS 467

RESULT 8
ID      003838      PRELIMINARY;      PRT;      465 AA.
AC      003838:
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      FILAGGRIN (PROFILAGGRIN) (Fragment).
GN      FLG.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RX      MEDLINE=91064347; PubMed=2248957;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 29:9432-9440(1990).
RN      (2)
RP      REVISIONS.
RX      MEDLINE=91255199; PubMed=2043621;
RA      Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
      gene.";
RL      Biochemistry 30:5814-5814(1991).
CC      -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
      PROMOTES DILIPID-BOND FORMATION AMONGST THE INTERMEDIATE
      FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
      -1- POLYMORPHISMS: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
      WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
      REPEATS.
      -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LONG,
      INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
      COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin
      GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
      PROTEOLYTICALLY CLEAVED.
      EMBL: M60499; AAA63246.1; -.
      DR      GO:0005198; F:structural molecule activity; IEA.
      DR      InterPro: IPR003303; Filaggrin.
      DR      Pfam: PF03516; Filaggrin; 3.
      DR      PRINTS: PR00487; FILAGGRIN.
      FT      NON_TER      1
      FT      NON_TER      1
      SQ      SEQUENCE 465 AA; 50280 MW; C983744C5E134097 CRC64;

Query Match      76.1%; Score 70; DB 4; Length 465;
Best Local Similarity 78.9%; Pred. No. 0.0052;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 SHOESTXGRSGRSGSGS 19
      ||||| |||||
DB      227 SHOESTXGRSGRSGSGS 245

RESULT 9
ID      075370      PRELIMINARY;      PRT;      322 AA.
AC      075370:
DT      01-NOV-1998 (TREMBlrel. 08, Created)
DT      01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT      01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE      Epidermal filaggrin (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99101527; PubMed=9886436;
RA      Girbal-Neuhausser E., Durieux J.U., Arnaud M., Dalbon P., Sebbag M.,
RA      Vincent C., Simon M., Senshu T., Masson-Bessiere C.,
RA      Jolivet-Reynaud C., Jolivet M., Seire G.;
RT      "The epitopes targeted by the rheumatoid arthritis-associated
      anti-filaggrin autoantibodies are posttranslationally generated on
      various sites of (pro)filaggrin by deamination of arginine residues.";
RL      U. Immunol. 162:585-594(1999).
DR      EMBL: AF043380; AAC23558.1; -.
DR      GO:0005198; F:structural molecule activity; IEA.
DR      InterPro: IPR003303; Filaggrin.
DR      Pfam: PF03516; Filaggrin; 2.
DR      PRINTS: PR00487; FILAGGRIN.
FT      NON_TER      1
FT      NON_TER      1
      SQ      SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match      75.0%; Score 69; DB 4; Length 322;
Best Local Similarity 77.8%; Pred. No. 0.0051;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 SHOESTXGRSGRSGSGS 18
      ||||| |||||
DB      305 SHOESTXGRSGRSGSGS 322

RESULT 10
ID      09P312      PRELIMINARY;      PRT;      822 AA.
AC      09P312:
DT      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Related to nucleolar phosphoprotein.
GN      B12P1.10.
OS      Neurospora crassa.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC      Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX      NCBI_TaxID=5141;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,
RA      Nyakatura G., Mewes H.W., Mannhaupt G.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN      (2)
RP      SEQUENCE FROM N.A.
RA      German Neurospora genome project;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AL390091; CAB98213.1; -.
DR      PIR: T51049; T51049.
DR      GO:0003676; F:nucleic acid binding; IEA.
DR      InterPro: IPR000504; RNA_rec_mot.
DR      InterPro: IPR005120; Smg-4_UPF3.
DR      Pfam: PF00076; rrm; 1.
DR      Pfam: PF03467; Smg4_UPF3; 1.
DR      SMART: SM00360; RRM; 1.
DR      PROSITE: PS50102; RRM; 1.
DR      PROSITE: PS00300; RRM_RNP_1; 1.
      SQ      SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match      59.8%; Score 55; DB 3; Length 822;
Best Local Similarity 58.8%; Pred. No. 2.3;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 QESTXGRSGRSGSGS 19
      ||||| |||||
DB      414 RESASGRTGRGRGCGT 430

RESULT 11

```

Q9V9Y3
ID Q9V9Y3 PRELIMINARY; PRT; 1284 AA.
AC Q9V9Y3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG11339-PA.
GN CG11339.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu B., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck U.,
RA Houtin D., Houston K.A., Howland T.J., Hernandez J.R., Houck U.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jimal B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li U., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.F., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheier F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Weissbach J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,
RA Zheng X.H., Zheng F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Beisman C.M., Bertan B.P., Carlson D.W., Celniker S.E.,
RA Clamp W.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smitnick F.,
RA Whitfield B.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA FlyBase;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003777; AAF57145.2; -
DR FlyBase; FBgn003841; CG11339.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_41; 1.
DR PRINTS; PRO035; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00577; FERM_3; 1.
SQ SEQUENCE 1284 AA; 13896 MW; EDP4119838A6447 CRC64;
Query Match 54.3%; Score 50; DB 5; Length 1284;
Best Local Similarity 61.1%; Pred. No. 23;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 SHQSTXGRSGRSG 18
DB 545 SDNESELSRSGRSG 562
RESULT 12
ID Q69375 PRELIMINARY; PRT; 820 AA.
AC Q69375;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative U25 protein.
OS Mouse cytomegalovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxId=10366;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K181;
RX MEDLINE=94233727; PubMed=7513920;
RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;
RT "Identification and characterization of a murine cytomegalovirus gene
RT with homology to the U25 open reading frame of human
RT cytomegalovirus.";
RL Virology 200; 643-650 (1994).
DR EMBL; U02500; AAA19449.1; -
DR InterPro; IPR006731; Herpes_pp85.
DR Pfam; PF04637; Herpes_pp85; 1.
SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;
Query Match 53.3%; Score 49; DB 12; Length 820;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 QESTXGRSGRS 14
DB 186 QSTXGRSGRS 197
RESULT 13
ID Q9CTM2 PRELIMINARY; PRT; 849 AA.
AC Q9CTM2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F1AC21.55
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv Columbia;
RX MEDLINE=21016719; PubMed=11130712;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-1
Perfect score: 92
Sequence: 1 SHQESTXGRSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	19	2	AAW61515
2	90	97.8	19	2	AAW61517
3	90	97.8	19	2	AAW61505
4	90	97.8	19	2	AAW61514
5	90	97.8	19	2	AAW61516
6	86	93.5	1467	5	ABB37805
7	84	91.3	19	2	AAW61506
8	84	91.3	19	2	AAW61508
9	84	91.3	19	2	AAW61507
10	84	91.3	19	2	AAW61512
11	84	91.3	19	2	AAW61509
12	84	91.3	19	2	AAW61511
13	84	91.3	19	2	AAW61513
14	84	91.3	19	2	AAW61510
15	81	88.0	21	2	AAW61520
16	81	88.0	22	4	AAW61520
17	59	64.1	330	2	AAW61520
18	59	64.1	330	2	AAW61520
19	59	64.1	330	2	AAW61520
20	55	59.8	330	2	AAW61520
21	51	55.4	1711	4	AAW78819
22	51	55.4	1711	4	AAW78819
23	50	54.3	477	6	ABO07142
24	50	54.3	1199	4	ABW58274
25	49	53.3	641	4	ABG19110

26	47	51.1	184	3	AAB57041	Aab57041 Human pro
27	46	50.0	125	2	AAV04858	Aay04858 Mycobacte
28	46	50.0	128	2	AAV04854	Aay04854 Mycobacte
29	46	50.0	443	4	AAW40069	Aam40069 Human pol
30	46	50.0	574	4	AAW41855	Aam41855 Human pol
31	46	50.0	700	4	AAW02421	Aae02421 Dario tel
32	46	50.0	752	4	AAU23535	Aau23535 Novel hum
33	45.5	49.5	135	4	ABW66181	Abb66181 Drosophi1
34	44	47.8	73	6	AAU54919	Aau54919 Propionib
35	44	47.8	73	6	AAU54919	AAU54919 Propionib
36	44	47.8	128	3	AAW33047	Aab33047 Pimus rad
37	44	47.8	229	7	AAW59825	Aae59825 Rat Prote
38	44	47.8	229	7	AAW59833	Aae59833 Rat Prote
39	44	47.8	229	7	AAW59837	Aae59837 Rat Prote
40	44	47.8	229	7	AAW59829	Aae59829 Rat Prote
41	44	47.8	356	5	AAW15535	AAW15535 Beta vulg
42	44	47.8	1038	7	AAW03412	AAW03412 Rice flow
43	43	46.7	92	4	AAU60843	Aau60843 Propionib
44	43	46.7	92	6	AAW57362	AAW57362 Propionib
45	43	46.7	106	3	AAW33285	Aag33285 Zea may

ALIGNMENTS

RESULT 1
ID AAW61515 standard; peptide, 19 AA.
AC AAW61515;
XX
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cFA, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure, Page 6, 19pp, English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
SQ Sequence 19 AA;

Query Match 98.8%; Score 91; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 3.5e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 2
 AAM61517 standard; peptide; 19 AA.

AC AAM61517;
 DT 26-OCT-1998 (first entry)
 XX Peptide cfo, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX MO9822503-A2.
 PN 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR 15-NOV-1996; 96NL-01004539.
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR Peptide derived from an antigen recognised by autoantibodies - is
 XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

QY Sequence 19 AA;
 SQ

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 5e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 3
 AAM61505 standard; peptide; 19 AA.

AC AAM61505;

XX 26-OCT-1998 (first entry)
 DT Peptide cfo1, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX MO9822503-A2.
 PN 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR Peptide derived from an antigen recognised by autoantibodies - is
 XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

QY Sequence 19 AA;
 SQ

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 5e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTXGRSRGRSGSGS 19
 DB 1 SHOESTXGRSRGRSGSGS 19

RESULT 4
 AAM61514 standard; peptide; 19 AA.

AC AAM61514;
 DT 26-OCT-1998 (first entry)
 XX Peptide cf, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.

```

XX  MO9822503-A2.
XX
XX  28-MAY-1998.
XX
XX  14-NOV-1997; 97MO-NL000624.
XX
XX  15-NOV-1996; 96NL-01004539.
XX
XX  (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX  (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX  Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX  WPI; 1998-398613/34.
XX
XX  Peptide derived from an antigen recognised by autoantibodies - is
XX  PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX  used in diagnosis of the disease.
XX
XX  Disclosure; Page 6, 19pp; English.
XX
XX  Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX  CC the profilaggrin antigen which is recognised by autoantibodies from
XX  CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX  CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX  CC The peptides were created by using standard solid phase synthesis, which
XX  CC produced them as peptide amides. These sequences may be used in the
XX  CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX  CC for obtaining polyclonal and monoclonal antibodies
XX
XX  Sequence 19 AA;
XX
XX  Query Match 97.8%; Score 90; DB 2; Length 19;
XX  Best Local Similarity 94.7%; Pred. No. 5e-07;
XX  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHOESTXGRSRGRSGRSGS 19
Db 1 SHOESTRGRSRGRSGRSGS 19
XX
RESULT 5
AAW61516
XX  AAW61516 standard; peptide; 19 AA.
XX
XX  AAW61516;
XX
XX  26-OCT-1998 (first entry)
XX
XX  Peptide cFE, based on cDNA of a profilaggrin repeat.
XX
XX  Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX  KM solid phase synthesis; peptide amide; polyclonal antibody;
XX  KM monoclonal antibody.
XX
XX  Synthetic.
XX  OS Homo sapiens.
XX
XX  MO9822503-A2.
XX
XX  28-MAY-1998.
XX
XX  14-NOV-1997; 97MO-NL000624.
XX
XX  15-NOV-1996; 96NL-01004539.
XX
XX  (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX  (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX  Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX  WPI; 1998-398613/34.
XX

```

```

XX  Peptide derived from an antigen recognised by autoantibodies - is
XX  PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX  used in diagnosis of the disease.
XX
XX  Disclosure; Page 6, 19pp; English.
XX
XX  Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX  CC the profilaggrin antigen which is recognised by autoantibodies from
XX  CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX  CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX  CC The peptides were created by using standard solid phase synthesis, which
XX  CC produced them as peptide amides. These sequences may be used in the
XX  CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX  CC for obtaining polyclonal and monoclonal antibodies
XX
XX  Sequence 19 AA;
XX
XX  Query Match 97.8%; Score 90; DB 2; Length 19;
XX  Best Local Similarity 94.7%; Pred. No. 5e-07;
XX  Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHOESTXGRSRGRSGRSGS 19
Db 1 SHOESTRGRSRGRSGRSGS 19
XX
RESULT 6
ABB97605
XX  ABB97605 standard; protein; 1467 AA.
XX
XX  ABB97605;
XX
XX  27-JUN-2002 (first entry)
XX
XX  Novel human protein SEQ ID NO: 873.
XX
XX  Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
XX  KM antinefertility; cerebroprotective; cytosclastic; rheumatic; gene therapy;
XX  KM neuroprotective; antiparkinsonian; protein therapy; EST;
XX  KM expressed sequence tag.
XX
XX  Homo sapiens.
XX  OS
XX  WO200222660-A2.
XX
XX  21-MAR-2002.
XX
XX  10-SEP-2001; 2001WO-US026015.
XX
XX  11-SEP-2000; 2000US-00659671.
XX
XX  (HSE-) HSESEQ INC.
XX
XX  Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX  PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX  DR WPI; 2002-292408/33.
XX  N-PSDB; AEN32791.
XX
XX  An isolated polynucleotide for treating diseases associated with its
XX  PT encoded polypeptide such as cancer and multiple sclerosis.
XX
XX  Example 2; SEQ ID NO 873; 509pp; English.
XX
XX  The present invention provides the protein and coding sequences of 444
XX  CC novel human proteins. These were isolated from expressed sequences tags
XX  CC (ESTs). They can be used to stimulate cell growth, to regulate
XX  CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX  CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX  CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX  CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX  CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

```


DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 11 /note= "Citruilline"
 XX
 PN MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY
 DB 1 SHOESTXGRSGRSGSGS 19
 1 SHOESTXGRSGRSGSGS 19
 RESULT 10
 AAM61512
 ID AAM61512 standard; peptide; 19 AA.
 XX
 AC AAM61512;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc3, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT
 PN MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY
 DB 1 SHOESTXGRSGRSGSGS 19
 1 SHOESTXGRSGRSGSGS 19
 RESULT 11
 AAM61509
 ID AAM61509 standard; peptide; 19 AA.
 XX
 AC AAM61509;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 16 /note= "Citruilline"
 XX
 PN MO9822503-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97MO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match          91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 SHQESTXGRSRGRSGSGS 19
DB 1 SHQESTXGRSRGRSGSGS 19

RESULT 12
AAW61511
ID AAW61511 standard; peptide; 19 AA.
XX
AC AAW61511;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 7 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
FT
FT
FT
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;
XX
Query Match          91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 4.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 SHQESTXGRSRGRSGSGS 19
DB 1 SHQESTXGRSRGRSGSGS 19

RESULT 13
AAW61513
ID AAW61513 standard; peptide; 19 AA.
XX
AC AAW61513;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc9, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 7 /note= "Citruilline"
FT Modified-site 16 /note= "Citruilline"
FT
FT
FT
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

```

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQSTXGSRGRSGSGS 19
 1 SHQSTXGSRGRSGSGS 19
 Db

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX
 DE Peptide cfc6, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX

OS Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 7 /note= "Citrulline"
 FT
 FT Modified-site 9 /note= "Citrulline"
 FT

XX WO9822503-A2.
 XX
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
 PI
 XX WPI, 1998-398613/34.
 DR

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHQSTXGSRGRSGSGS 19
 1 SHQSTXGSRGRSGSGS 19
 Db

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX
 XX Peptide XI based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX

OS Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 9 /note= "Citrulline"
 FT
 FT Modified-site 9 /note= "Citrulline"
 FT

XX WO9822503-A2.
 XX
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
 PI
 XX WPI, 1998-398613/34.
 DR

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 21 AA;

Query Match 88.0%; Score 81; DB 2; Length 21;
 Best Local Similarity 94.4%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HQESTXGSRGRSGSGS 19
 4 HQESTXGSRGRSGSGS 21
 Db

Search completed: September 28, 2004, 06:24:09
 Job time : 57.1302 secs

Blank Sheet

Tue Sep 28 07:33:33 2004

us-09-308-150-1.open.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-1

Perfect score: 92

Sequence: 1 SHOESTXGRSGRSGSGS 19

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	19	9	US-09-308-150-12
2	90	97.8	19	9	US-09-308-150-1
3	90	97.8	19	9	US-09-308-150-11
4	90	97.8	19	9	US-09-308-150-13
5	90	97.8	19	9	US-09-308-150-14
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-3
8	84	91.3	19	9	US-09-308-150-4
9	84	91.3	19	9	US-09-308-150-5
10	84	91.3	19	9	US-09-308-150-6
11	84	91.3	19	9	US-09-308-150-7
12	84	91.3	19	9	US-09-308-150-8
13	84	91.3	19	9	US-09-308-150-9
14	81	88.0	21	9	US-09-308-150-10
15	81	88.0	22	9	US-09-747-0239A-22

15	50	54.3	477	15	US-10-161-927-62	Sequence 62, Appl
17	50	54.3	1087	9	US-09-918-909-24	Sequence 24, Appl
18	50	54.3	1087	16	US-10-641-991-24	Sequence 24, Appl
19	48	52.2	179	12	US-10-425-114-53545	Sequence 53545, A
20	48	52.2	436	15	US-10-369-493-8178	Sequence 8178, Ap
21	48	52.2	506	16	US-10-437-963-185974	Sequence 185974, A
22	47	51.1	50	14	US-10-029-386-27692	Sequence 27692, A
23	47	51.1	123	16	US-10-767-701-32436	Sequence 32436, A
24	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
25	46	50.0	164	16	US-10-437-963-112419	Sequence 112419, A
26	46	50.0	384	12	US-10-425-114-72136	Sequence 72136, A
27	46	50.0	384	12	US-10-425-114-72137	Sequence 72137, A
28	46	50.0	436	14	US-10-156-761-13022	Sequence 13022, A
29	46	50.0	580	12	US-10-425-114-70952	Sequence 70952, A
30	46	50.0	612	12	US-10-424-599-230819	Sequence 230819, A
31	45	48.9	223	16	US-10-437-963-133033	Sequence 133033, A
32	45	48.9	838	16	US-10-437-963-146503	Sequence 146503, A
33	45	48.9	854	12	US-10-425-114-57838	Sequence 57838, A
34	44.5	48.4	85	12	US-10-424-559-161928	Sequence 161928, A
35	44	47.8	52	16	US-10-437-963-155141	Sequence 155141, A
36	44	47.8	76	12	US-10-424-599-248912	Sequence 248912, A
37	44	47.8	78	16	US-10-437-963-203073	Sequence 203073, A
38	44	47.8	105	16	US-10-437-963-196208	Sequence 196208, A
39	44	47.8	131	16	US-10-767-701-33185	Sequence 33185, A
40	44	47.8	232	16	US-10-437-963-183394	Sequence 183394, A
41	44	47.8	262	12	US-10-425-114-72166	Sequence 72166, A
42	44	47.8	262	12	US-10-425-114-72167	Sequence 72167, A
43	44	47.8	340	12	US-10-425-114-68797	Sequence 68797, A
44	44	47.8	406	12	US-10-425-114-41013	Sequence 41013, A
45	44	47.8	843	14	US-10-359-431-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
TITLE OF INVENTION: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY ANTIBODIES DERIVED FROM PATIENTS WITH RHEUMATOID ARTHRITIS, TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 98.9%; Score 91; DB 9; Length 19;
Best Match Similarity 94.7%; Pred No. 2.9e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHOESTXGRSGRSGSGS 19

Db 1 SHOESTXGRSRGRSGSGS 19

RESULT 2

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; OTHER INFORMATION: Xaa is citrulline
 ; US-09-308-150-1

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGSGS 19
 Db 1 SHOESTXGRSRGRSGSGS 19

RESULT 3

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4e-06;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGSGS 19
 Db 1 SHOESTXGRSRGRSGSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 ; US-09-308-150-13

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSRGRSGSGS 19
 Db 1 SHOESTXGRSRGRSGSGS 19

RESULT 5

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 19

```
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 14
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
US-09-308-150-14
OTHER INFORMATION: known cDNA sequences of human profilaggrin
```

```
Query Match      97.8% Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 6
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-2
OTHER INFORMATION: Xaa is citrulline
```

```
Query Match      91.3% Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 7
US-09-308-150-3
Sequence 3, Application US/09308150
```

```
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-3
OTHER INFORMATION: Xaa is citrulline
```

```
Query Match      91.3% Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTXGRSGRSGSGS 19
Db      1 SHOESTXGRSGRSGSGS 19
```

```
RESULT 8
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-4
OTHER INFORMATION: Xaa is citrulline
```

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 9
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 10
US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 11
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTXGRSGRSGSGS 19

RESULT 12
US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1  SHOESTXGRSRGRSGSGS 19
DB      1  SHOESTXGRSRGRSGSGS 19

```

```

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1  SHOESTXGRSRGRSGSGS 19
DB      1  SHOESTXGRSRGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match          88.0%; Score 81; DB 9; Length 21;
Best Local Similarity 94.4%; Pred. No. 8.3e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2  HOESTXGRSRGRSGSGS 19
DB      4  HOESTXGRSRGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.003INPUS00 INNS:031
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US/09/747,029A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```


ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25967

Query Match 51.1%; Score 47; DB 4; Length 421;
Best Local Similarity 52.9%; Pred. No. 9.8;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGRSG 18
DB 13 HARTGTGTGRCGRSG 29

RESULT 3
US-09-489-039A-7836
Sequence 7836, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489.039A

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7836

LENGTH: 491

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-7836

Query Match 48.9%; Score 45; DB 4; Length 491;
Best Local Similarity 69.2%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 STXGRSGRSGRSG 17
DB 217 TAGRSGRSGRSG 229

RESULT 4
US-09-252-991A-25967
Sequence 25967, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 25967

LENGTH: 518

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25967

Query Match 47.8%; Score 44; DB 4; Length 518;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 STXGRSGRSGRSG 18
DB 412 SGGGRSGRSGRSG 425

RESULT 5

US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T

NUMBER OF SEQUENCES: 99
Lymphocyte Responses to Hepatitis B Virus

RESPONSE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/591.502B

FILING DATE: 20-May-1996

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,540

FILING DATE: 26-AUG-1991

APPLICATION NUMBER: US 07/935,898

FILING DATE: 26-AUG-1992

APPLICATION NUMBER: US 08/100,870

FILING DATE: 02-AUG-1993

APPLICATION NUMBER: WO PCT/US94/08685

FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Ellen Lauver

REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 014740-000230US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 843 amino acids

TYPE: amino acid

STRANDEDNESS: <unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 45:

US-08-591-502B-45

Query Match 47.8%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 57;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGRSG 19
DB 226 QPQGSYARXGRSGRSG 242

RESULT 6
US-09-252-991A-17002
Sequence 17002, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 17002
;; LENGTH: 326
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17002

Query Match 47.3%; Score 43.5; DB 4; Length 326;
Best Local Similarity 68.8%; Pred. No. 26;
Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 4 ESTXGRSGRSGSGS 19
Db 236 ESSSGSRARS-RDGS 250

RESULT 7
US-09-252-991A-28985
; Sequence 28985, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28985
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28985

Query Match 46.7%; Score 43; DB 4; Length 239;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SHQSTXGRSGRSGRS 17
Db 171 SHDRAPGRSRDRAGKA 187

RESULT 8
US-09-252-991A-21702
; Sequence 21702, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21702
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21702

Query Match 46.7%; Score 43; DB 4; Length 395;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHQSTXGRSGRSGRS 18
Db 251 SHQASGRGDPHESGRPG 268

RESULT 9
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015

Query Match 46.7%; Score 43; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 HQESTXGRSGRSGRS 18
Db 94 HAEQDPGAGTGRAGRAG 110

RESULT 10
US-09-252-991A-24620
; Sequence 24620, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24620
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24620

Query Match 46.7%; Score 43; DB 4; Length 546;
Best Local Similarity 64.3%; Pred. No. 52;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 STXGRSGRSGRS 18
Db 427 SRLGRDRGASGRPG 440

RESULT 11
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:

Query Match	46.7%;	Score 43;	DB 4;	length 724;
Best local Similarity	60.0%;	Pred. No. 69;		

RESULT 15
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

```

; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hoegson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523, 849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO: 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36

```

```

Query Match          46.7%; Score 43; DB 4; Length 745;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      5 STXGRSRGRGRSGS 19
Db      585 SPGHSQGRGRGASGS 599

```

Search completed: September 28, 2004, 06:26:40
Job time : 15.4479 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-2

Sequence: 1 SHQESTRGXSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	97.8	2248	2	A35938	profilaggrin - hum
2	84	91.3	416	2	A32947	filaggrin precursor
3	75	81.5	591	2	A45135	profilaggrin - hum
4	58	63.0	506	1	W2WL47	E2 protein - human
5	50	54.3	800	2	T02852	probable membrane
6	48	52.2	822	2	T51049	related to nucleol
7	47	51.1	471	2	T33997	hypothetical prote
8	46	50.0	197	2	T18918	hypothetical prote
9	46	50.0	203	2	C87801	protein C10G1.9.1
10	46	50.0	399	2	T46259	hypothetical prote
11	46	50.0	836	2	G84727	probable DNA topoi
12	45	48.9	374	2	T33328	hypothetical prote
13	45	48.9	1337	2	T30291	hypothetical prote
14	45	48.9	1791	2	T03245	hypothetical prote
15	44	47.8	229	2	UC7219	nuclear protein SR
16	44	47.8	790	2	T05576	hypothetical prote
17	44	47.8	1232	2	S40766	hypothetical prote
18	43.5	47.3	463	2	T51194	hypothetical prote
19	43	46.7	123	2	T16234	hypothetical prote
20	43	46.7	409	2	T35118	probable secreted
21	43	46.7	452	2	S35482	E2 protein - human
22	43	46.7	525	2	T53194	DNA-directed DNA p
23	43	46.7	827	1	T13468	DNA-directed DNA p
24	43	46.7	842	1	JDVLVS	DNA-directed DNA p
25	43	46.7	843	1	JDVLVR	hypothetical prote
26	42	45.7	203	2	T25916	hypothetical prote
27	42	45.7	306	2	T21220	wingless receptor
28	42	45.7	694	2	S71785	hypothetical prote
29	42	45.7	849	2	A96592	hypothetical prote

30	42	45.7	889	2	T33422	hypothetical prote
31	42	45.7	1287	2	S55954	viral mRNA transla
32	42	45.7	1829	2	T35681	probable sensory h
33	42	45.7	2233	2	S63347	acetyl-CoA carboxy
34	41.5	45.1	646	2	H96665	protein F2C12.10
35	41.5	45.1	1895	2	T06609	disease resistance
36	41	44.6	136	2	T35632	probable transposa
37	41	44.6	200	1	OCB827	BDLP protein - hu
38	41	44.6	205	2	A19634	nucleolar protein
39	41	44.6	312	2	A31846	130K paracrystalli
40	41	44.6	315	2	S53589	SOL2 protein - yea
41	41	44.6	360	2	G71444	hypothetical prote
42	41	44.6	373	2	T02976	probable DNA bindi
43	41	44.6	408	2	T21412	hypothetical prote
44	41	44.6	458	2	G83690	hypothetical prote
45	41	44.6	461	2	S36593	E2 protein - human

ALIGNMENTS

```
RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GNN>
A:Cross-references: GB:J02929
A:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
P:245-569/Region: filaggrin repeat
F:570-893/Region: filaggrin repeat
F:1074-1397/Region: filaggrin repeat
F:1573-1896/Region: filaggrin repeat

Query Match          97.8%; Score 90; DB 2; Length 2248;
Best Local Similarity 94.7%; Pred. No. 8.9e+06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 SHQESTRGXSRGSRGSGS 19
DB      551 SHQESTRGXSRGSRGSGS 569

RESULT 2
A32947
filaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cammizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A:Title: Characterization of a cDNA clone encoding human filaggrin and localization of th
A:Reference number: A32947; MUID:99296901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M2435; NID:g182604; PIND:AA52454.1; PID:g182605
A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:
```

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 91.3%; Score 84; DB 2; Length 416;
 Best Local Similarity 89.5%; Pred. No. 1.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 |||||
 DB 7 SHQSTRGXSRGSRGSGS 25

RESULT 3
 A45135
 A:Protein: profilaggrin - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135
 R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nixunskisiri, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992
 A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717
 A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101089; GB:M90967; NID:q190408; PIDN:AAA60177.1; PID:9553621
 A:Note: sequence extracted from NCBI backbone (NCBI:P:118773)
 C:Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 81.5%; Score 75; DB 2; Length 591;
 Best Local Similarity 84.2%; Pred. No. 0.00051;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 |||||
 DB 449 SHQSTRGXSRGSRGSGS 467

RESULT 4
 W2ML47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D5324
 R:Kiyono, T.; Adachi, A.; Ishibashi, M.
 Virology 177, 401-405, 1990
 A:Title: Genome organization and taxonomic position of human papillomavirus type 47
 A:Reference number: A5324; MUID:90281611; PMID:2122112
 A:Accession: D5324
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KTY>
 A:Cross-references: GB:M2305; NID:g333062; PIDN:AAA46979.1; PID:g333067
 C:Superfamily: papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.19;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHQSTRGXSRGSRGSGS 19
 |::||| |||::||

DB 342 SREGNTRGGRGGRGGRGSGS 360

RESULT 5
 T02852
 Probable membrane protein U1439.4 [imported] - Leishmania major (strain Friedlin)
 C:Species: Leishmania major
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C:Accession: T02852; H81462
 R:Myler, P.J.
 submitted to the EMBL Data Library, May 1998
 A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
 A:Reference number: Z14740
 A:Accession: T02852
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-800 <MYL>
 A:Cross-references: EMBL:AE001274; NID:g3264850; PID:g2266920
 R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; f
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A:Reference number: A81455; MUID:99178987; PMID:10077609
 A:Accession: H81462
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-800 <PYL>
 A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AA024675.1; PID:g2266920; GSPDB:GN001
 A:Experimental source: strain MEOM/IL/81/Friedlin
 C:Genetics:
 A:Gene: U1439.4
 A:Map position: 1

Query Match 54.3%; Score 50; DB 2; Length 800;
 Best Local Similarity 55.6%; Pred. No. 5.2;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 HQSTRGXSRGSRGSGS 19
 ||::|||
 DB 429 HRDGVGRGSRGSRGSGS 446

RESULT 6
 T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12FL10
 C:Species: Neurospora crassa
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049
 R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: Z25286
 A:Accession: T51049
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12FL1.10
 A:Experimental source: BAC clone B12FL1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B12FL1.10
 A:Map position: 6
 A:introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 11;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTRGXSRGSRGSGS 19
 |||||
 DB 414 RESAGRTGRGRGRCGT 430

RESULT 7
 T33997

C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Sep-2000
 C/Accession: T33328
 R:Nohlmann, P.; Hawkins, J.; Gillam, B.
 Submitted to the EMBL Data Library, July 1998
 A/Description: The sequence of C. elegans cosmid F13C5.
 A/Reference number: Z21324
 A/Accession: T33328
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-374 <NOH>
 A/Cross-references: EMBL:AB077531; PIDN:AA04610.1; GSPDB:GN00028; CESP:F13C5.2
 A/Experimental source: strain Bristol N2; clone F13C5
 C/Genetics:
 A/Gene: CESP:F13C5.2
 A/Map position: X
 A/Intons: 25/3; 135/3; 189/3; 313/1
 C/Superfamily: bromodomain homology
 F141-198/Domain: bromodomain homology <BRO>

Query Match 48.9%; Score 45; DB 2; Length 374;
 Best Local Similarity 69.2%; Pred. No. 15;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ESTRGXSRGSGR 16
 DB 64 ESTRGCTGSGR 76

RESULT 13
 T30291
 dextranase - Streptococcus sobrinus
 C/Species: Streptococcus sobrinus
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30291
 R:Wanda, S.Y.; Curtiss III, R.
 J. Bacteriol. 176, 3839-3850, 1994
 A/Title: Purification and characterization of Streptococcus sobrinus dextranase produced
 A/Reference number: Z20810; MUID:94292401; PMID:80211165
 A/Accession: T30291
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1337 <MAN>
 A/Cross-references: EMBL:M96978; NID:9450640; PID:9450641; PIDN:AAA21772.1

Query Match 48.9%; Score 45; DB 2; Length 1337;
 Best Local Similarity 47.1%; Pred. No. 51;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTRGXSRGSGRS 17
 DB 1154 ANQDSTKGSSADQSGKS 1170

RESULT 14
 T02345
 hypothetical protein KIAA0324 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
 C/Accession: T02345
 R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, O.; White, S.; Deng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 submitted to the EMBL Data Library, March 1998
 A/Description: Sequencing of human chromosome 16p13.3.
 A/Reference number: Z14664
 A/Accession: T02345
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1791 <RIC>
 A/Cross-references: EMBL:AC004493; NID:92996648; PIDN:AA008453.1; PID:92996650
 C/Genetics:
 A/Map position: 16
 A/Intons: 1610/2; 1706/2

A/Note: KIAA0324

Query Match 48.9%; Score 45; DB 2; Length 1791;
 Best Local Similarity 58.8%; Pred. No. 68;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTRGXSRGSGRS 17
 DB 858 AROESSRSTSSRRRGRS 874

RESULT 15
 UC7219
 nuclear protein SR-25 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
 C/Accession: JC7219
 R:Sasahara, K.; Yamaoka, T.; Moritani, M.; Tanaka, M.; Iwahana, H.; Yoshimoto, K.; Miyase
 Biochem. Biophys. Res. Commun. 269, 444-450, 2000
 A/Title: Molecular cloning and expression analysis of a putative nuclear protein, SR-25.
 A/Reference number: JC7219; MUID:20175222; PMID:10708573
 A/Accession: UC7219
 A/Molecule type: mRNA
 A/Residues: 1-229 <SAS>
 A/Cross-references: DBJ:AB035383; NID:97619895; PIDN:BAA94743.1; PID:97619896
 A/Experimental source: MIN6 cell line
 A/Comment: This protein is a highly hydrophilic nuclear protein with a serine-arginine re
 A/splicing factors
 A/keywords: nucleus; RNA processing

Query Match 47.8%; Score 44; DB 2; Length 229;
 Best Local Similarity 52.9%; Pred. No. 13;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 CESTRGXSRGSGRS 19
 DB 7 RKRSRSGRSRSGRS 23

Search completed: September 28, 2004, 06:14:57
 Job time : 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHQESTRGXSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	91.3	416	1	FILA_HUMAN
2	58	63.0	506	1	VE2_HPV47
3	45	48.9	1337	1	DEXT_STRDO
4	44	47.8	1232	1	YQO5_CAEEL
5	43	46.7	123	1	LSM4_CAEEL
6	43	46.7	452	1	VE2_HPV17
7	43	46.7	525	1	MRP1_RAT
8	43	46.7	730	1	DPO1_HPBV4
9	43	46.7	842	1	DPO1_HPBV4
10	43	46.7	843	1	DPO1_HPBV4
11	42.5	46.2	388	1	ROG_MOUSE
12	42.5	46.2	391	1	ROG_HUMAN
13	42.5	46.2	1180	1	TYK2_MOUSE
14	42	45.7	694	1	FRZ2_DROME
15	42	45.7	1097	1	RNT1_FUGRU
16	42	45.7	1287	1	SKT2_YEAST
17	42	45.7	2233	1	COAC_YEAST
18	41.5	45.1	1895	1	WR19_ARATH
19	41	44.6	197	1	RX21_DROME
20	41	44.6	205	1	GAR1_YEAST
21	41	44.6	315	1	SOL2_YEAST
22	41	44.6	461	1	VE2_HPV09
23	41	44.6	503	1	VE2_HPV21
24	41	44.6	772	1	MR11_CAEEL
25	41	44.6	774	1	MRK2_MOUSE
26	41	44.6	825	1	SES_RAT
27	41	44.6	955	1	CLIS2_HUMAN
28	41	44.6	2404	1	SON_MOUSE
29	41	44.6	2426	1	SON_HUMAN
30	40.5	44.0	255	1	THO4_MOUSE
31	40.5	44.0	257	1	THO4_HUMAN
32	40.5	44.0	1509	1	MYSN_ACACA
33	40	43.5	414	1	NSR1_YEAST

ALIGNMENTS

RESULT 1	ID	FILA_HUMAN	STANDARD;	PRT;	416 AA.
AC	P20930				
DT	01-FEB-1991	(Rel. 17, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
D3	Filaggrin precursor	(Fragment).			
GN	FLG				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=89296901; PubMed=2740331;				
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,				
RA	Canizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.,				
RT	"Characterization of a cDNA clone encoding human filaggrin and				
RT	localization of the gene to chromosome region 1q21."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).				
RM	[2]				
RM	CITRULLINATION.				
RX	MEDLINE=96374388; PubMed=8780679;				
RA	Senshu T., Kan S., Ogawa H., Manabe M., Asaga H.,				
RT	"Preferential delamination of keratin K1 and filaggrin during the				
RT	terminal differentiation of human epidermis."				
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).				
CC	-I- FUNCTION: Aggregates keratin intermediate filaments and promotes				
CC	disulfide-bond formation among the intermediate filaments during				
CC	terminal differentiation of mammalian epidermis.				
CC	-I- PFM: Filaggrin is initially synthesized as a large, insoluble,				
CC	highly phosphorylated precursor containing many tandem copies of				
CC	324 AA, which are not separated by "large linker". The precursor				
CC	is deposited as keratohyalin granules. During terminal				
CC	differentiation it is dephosphorylated and proteolytically				
CC	cleaved.				
CC	-I- PFM: Undergoes delamination of some arginine residues				
CC	(citrullination).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M24355; AAA52454.1; -				P54672 dictyostell
DR	PIR; A32947; A32947.				P36782 human papil
DR	Gene; HGNC:3748; FLG.				Q81361 mus musculu
DR	MIM; 135940; -				P22807 drosophila
DR	GO; GO:0005882; C:intermediate filament; NAS.				Q00192 escherichia
DR	GO; GO:0005198; F:structural molecule activity; NAS.				Q00586 crepusculi
DR	GO; GO:0007275; P:development; NAS.				P02463 mus musculu
DR	InterPro; IPR003303; Filaggrin.				Q16629 homo sapien

```

DR Pfam; PF03516; Filaggrin; 2.
DR PRINTS; PR00487; FILAGRIN.
KM Phosphorylation; Citrullination; Developmental protein.
FT NON TER 1
SQ SEQUENCE 416 AA; 44105 MW; DEBA3218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
Best Local Similarity 89.5%; Pred. No. 5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRCKSGRSGRSGS 19
Db 7 SHOESTRCKSGRSGRSGS 25

RESULT 2
VE2 HPV47 STANDARD; PRT; 506 AA.
ID VE2 HPV47
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human Papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
RT type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2E RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2E'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33305; AAA46979.1; -.
DR PIR; D35324; W2ML47.
DR HSP; P03122; ZHOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR Pfam; PF00511; E2_C; 1.
DR Pfam; PF00508; E2_N; 1.
DR ProDom; PD000672; E2_N; 1.
DR ProDom; PD000678; E2_N; 1.
KM Early protein; Transcription regulation; Activator; DNA-binding;
KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF825065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.076;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOESTRCKSGRSGRSGS 19

```

```

Db 342 SRENTGRGRGRGRGRGS 360

RESULT 3
DEX STRDO STANDARD; PRT; 1337 AA.
ID DEX STRDO
AC P39653;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
DE glucanohydrolase).
GN DEX.
OS Streptococcus downei (Streptococcus sobrinus).
OG Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RX SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
RX STRAIN=6715 / UNB66;
RX MEDLINE=94292401; PubMed=8021165;
RA Wanda S.-Y., Cutliss R. Iii;
RT "Purification and characterization of Streptococcus sobrinus
RT dextranase produced in recombinant Escherichia coli and sequence
RT analysis of the dextranase gene.";
RL Bacteriol. 176:3839-3850(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
CC PELVICLE-COATED TOOTH SURFACE.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
CC linkages in dextran.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
CC and at 35 degrees Celsius.
CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M56978; AAA21772.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfams; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KM Hydrolyase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
KM Signal; Plasmid.
FT SIGNAL 1 30
FT CHAIN 31 1308 DEXTRANASE.
FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 24;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRCKSGRSGRSGS 17
Db 1154 ANQDSTKSSADQSGKS 1170

RESULT 4
Y005_CAEEL

```

```

ID Y005 CAEBL STANDARD; PRT; 1232 AA.
AC P34643;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK512.5 in chromosome III.
GN ZK512.5.
OS Caenorhabditis elegans.
OC Rhabditidae; Pelecerozoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelecerozoa; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Burroughs J., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copestake T., Cooper J., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirwan J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rinken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaiden N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaughan M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.,
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RT Nature 368:32-38(1994).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z22177; CAA80146.1; -
CC PIR; S40766; S40766.
CC WormRep; ZK512.5; CE00411.
CC Hypothetical protein.
CC SEQUENCE 1232 AA; 134923 MW; 6DFC35D664A8D6A CRC64;
SQ
Query Match 47.8%; Score 44; DB 1; Length 1232;
Best Local Similarity 64.3%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 HOESTRGXRGGRSG 15
DB 110 HQNSRGPSRGSG 123
RESULT 5
LSM4 CAEBL STANDARD; PRT; 123 AA.
ID LSM4 CAEBL
AC 01952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein Lsm4.
GN F32A5.7.
OS Caenorhabditis elegans.
OC Rhabditidae; Pelecerozoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelecerozoa; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Paulley A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (by similarity).

```

```

CC -!- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (by
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U20864; AAC46661.1; -
CC PIR; T16234; T16234.
CC WormRep; F32A5.7; CE01277.
CC InterPro; IPR006649; snRNP.
CC InterPro; IPR01163; snRNP_Sm.
CC Pfam; PF01423; LSM; 1.
CC ProDom; PD020287; snRNP; 1.
CC SMART; SM0051; Sm; 1.
CC Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
CC RNA-binding.
CC SEQUENCE 123 AA; 13593 MW; B3E874B67705109 CRC64;
SQ
Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 3; 7;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QESTRGXRGGRSG 18
DB 93 REGSGRGGRSG 108
RESULT 6
ID V02 HPV17 STANDARD; PRT; 452 AA.
AC P36785;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2
OS Human papillomavirus type 17.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10607;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.,
RA "Primer-directed sequencing of human papillomavirus types.",
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -!- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNNGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -!- SUBUNIT: Binds DNA as a dimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74469; CAA52515.1; -

```



```

DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P
OS Hepatitis B virus (subtype adr / mutant).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31512;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89183619; PubMed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
RT virus subtype adr."/
RL Nucleic Acids Res. 17:2124-2124(1989).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14193; CAA32399.1; -.
DR EMBL; X14193; CAA32405.1; ALT_TERM.
DR PIR; S04568; JDIVIS.
DR InterPro; IPR001462; DNAPOL_viral_C.
DR InterPro; IPR000201; DNAPOL_viral_N.
DR InterPro; IPR000477; RVise.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvc; 2.
DR ProDom; PD000814; DNAPOL_viral_C; 1.
KW Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 842 AA; 94545 MW; 137FE37246BA5C6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 842;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242
:|:|:|:|:|
:|:|:|:|:|

RESULT 10
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID DPOL_HPBVR
AC P03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE p protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=31512;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada K., Igatahi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adr."/
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA) (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00867; -. NOT_ANNOTATED_CDS.
DR PIR; A00704; JDIVLR.
DR InterPro; IPR001462; DNAPOL_viral_C.
DR InterPro; IPR000201; DNAPOL_viral_N.
DR InterPro; IPR000477; RVise.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; rvc; 2.
DR ProDom; PD000814; DNAPOL_viral_C; 1.
KW Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C4E8B CRC64;

Query Match 46.7%; Score 43; DB 1; Length 843;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
DB 233 ARKSGRSGS 242
:|:|:|:|:|
:|:|:|:|:|

RESULT 11
ROG_MOUSE STANDARD; PRT; 388 AA.
ID ROG_MOUSE
AC O35479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein G (hnRNP G) (RNA binding motif
DE protein, X chromosome).
DE RBMX OR RBMX1 OR RBMX2 OR HNRPG OR HNRPG.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98119027; PubMed=9457683;
RA Delbridge M.L., Ma K., Subbarao M.N., Cooke H.J., Basan S.;
RT "Evolution of mammalian HNRPG and its relationship with the putative
RT azoospermia factor RBM."/
RL Mamm. Genome 9:168-170(1998).
CC -1- FUNCTION: Binds RNA.
CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes.
CC -1- PTM: O-glycosylated (by similarity).
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031568; AAB86639.1; -.
DR HSSP; P09651; 1HA1.
DR MG; MG1:1343045; Rbmxt.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.

```

DR PROSITE PS00030; RRM RNP 1; 1.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Glycoprotein.
FT DOMAIN 8 86 RNA-BINDING (RRM).
SQ SEQUENCE 388 AA; 4223 MW; F02FF2DCBCC87F14 CXC64;
Query Match 46.2%; Score 42.5; DB 1; Length 388;
Best Local Similarity 52.6%; Pred. No. 16;
Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
QY 1 SHGSTRGKSGRGSRGSGS 19
Db 360 SYSSSSRGAPRG-GGRGGS 377
RESULT 12
ROG_HUMAN
ID ROG_HUMAN STANDARD; PRT; 391 AA.
AC P38135; Q96933;
DC 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Heterogeneous nuclear ribonucleoprotein G (hnRNP G) (RNA binding motif
DE protein, X chromosome) (Glycoprotein p43).
GN BMX OR BMXPI OR HNRPG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
CX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=94021365; PubMed=7692398;
RA Soulaad M., della Valle V., Slomi M., Pinol-Roma S., Codogno P.,
RA Bavy C., Belli M., Lacroix J.-C., Monod G., Dreyfuss G.,
RA Larsen C.-J.;
RT "hnRNP G: sequence and characterization of a glycosylated RNA-binding
RL protein.";
RL Nucleic Acids Res. 21:4210-4217(1993).
RN [2]
RP REVISIONS.
RA Venables J.P., Larsen C.-J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lingenteleer P.A., Delbridge M.L., Thomas S., Graves J.A.,
RA Distech C.M.;
RT "Expression and conservation of processed copies of the BMX gene.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepien M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raza S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gnatratie P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
-1- FUNCTION: Binds RNA.

```

CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes.
CC -1- PM: O-glycosylated.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; Z23064; CAAB059.1; -.
DR EMBL; AF266723; AAK58567.1; -.
DR EMBL; AF266720; AAK58567.1; JOINED.
DR EMBL; AF266721; AAK58567.1; JOINED.
DR EMBL; AF266722; AAK58567.1; JOINED.
DR EMBL; BC006550; AAH0550.1; -.
DR EMBL; BC007435; AAH07435.1; -.
DR HSSP; P09651; IHAL.
DR SWISS-2DPAGE; P38159; HUMAN.
DR Genew; HGNC:9910; RBMX.
DR GK; P38159; -.
DR MIM; 300199; -.
DR GO; GO:003050; C:heterogeneous nuclear ribonucleoprotein com. . .; NAS.
DR GO; GO:0033723; F:RNA binding; NAS.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
DR Nuclear protein; RNA-binding; Ribonucleoprotein; Glycoprotein.
FT DOMAIN 8 86 RNA-BINDING (RRM).
FT CONFLOC 259 259 G -> E (IN REF. 1).
SQ SEQUENCE 391 AA; 42331 MW; 904FEB9BFC573546 CRC64;

Query March 46.2%; Score 42.5; DB 1; Length 391;
Best local similarity 52.6%; Pred. No. 16;
Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 SHOESTRGXSGRSGSGS 19
Db 363 SYSSSSRGAPRG-GGRGGS 380

RESULT 13
TYK2_MOUSE STANDARD; PRT; 1180 AA.
ID TYK2_MOUSE
AC Q9R117; O88431; Q8VE41;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Non-receptor tyrosine-protein kinase TYK2 (EC 2.7.1.112).
GN TYK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RX [1]
RX SEQUENCE FROM N.A.
RP STRAIN=129/Sv; TISSUE=liver;
RX MEDLINE=21000430; PubMed=11070173;
RA Karaghiosoff M., Neubauer H., Laessing C., Kovarik P., Schindler H.,
RA Pitcher H., McCoy B., Bogdan C., Decker T., Brem G., Pfeiffer K.,
RA Meiller M.;
RT "Partial impairment of cytokine responses in Tyk2-deficient mice.";
RL Immunity 13:549-560(2000).
RX [2]
RX SEQUENCE OF 769-1180 FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushih T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton S., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Involved in intracellular signal transduction by
 CC amplifying type I and type II IFN signaling. Phosphorylates the
 CC interferon-alpha/beta receptor alpha chain. Plays an essential
 CC role in promoting selective immune responses, including innate
 CC host defense mechanisms and specific antiviral activities.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC MISCELLANEOUS: Tyk2 deficient mice display reduced responses to
 CC interferon-alpha/beta and IL-12 and a selective deficiency in
 CC Stat3 activation in these pathways. Interferon-gamma signaling is
 CC also impaired.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. JAK
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 FERM domain.
 CC -1- SIMILARITY: Contains 1 SH2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF173032; AAD49423.1; -;
 DR EMBL: AF052607; AAC34580.2; -;
 DR EMBL: BC019789; AAI19789.1; -;
 DR HSSP: P1362; IFGK.
 DR MGI: 1929470; Tyk2.
 DR InterPro: IPR000299; Band 4.1.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_Ehr_kinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR008266; Tyr_kinase_AS.
 DR Pfam: PF00069; Pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Prot_kinase; 2.
 DR SMART: SM00295; B4; 1.
 DR SMART: SM00220; S_TKc; 2.
 DR SMART: SM00219; TYRKc; 2.
 DR PROSITE: PS00660; FERM_1; FALSE_NEG.
 DR PROSITE: PS00661; FERM_2; FALSE_NEG.
 DR PROSITE: PS00577; FERM_3; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_DOM; 2.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00001; SH2; FALSE_NEG.
 DR Trnasease: Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW SH2 domain; Repeat.
 FT DOMAIN 29 426 FERM.
 FT DOMAIN 445 525 SH2 (ATYPICAL).
 FT DOMAIN 585 862 PROTEIN KINASE 1.
 FT DOMAIN 890 1162 PROTEIN KINASE 2.
 FT NP_BIND 896 904 ATP (BY SIMILARITY).
 FT BINDING 923 923 ATP (BY SIMILARITY).

FT ACT SITE 1016 1016 BY SIMILARITY.
 FT MOD_RES 1047 1047 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 191 191 C -> F (IN REF. 1; AAC34580).
 FT CONFLICT 234 234 H -> R (IN REF. 1; AAC34580).
 FT CONFLICT 355 355 K -> E (IN REF. 1; AAC34580).
 FT CONFLICT 831 831 C -> S (IN REF. 1).
 FT CONFLICT 1099 1099 T -> M (IN REF. 1).
 SQ SEQUENCE 1180 AA; 132857 MW; 65E060079F08B1C9 CRC64;
 Query Match 46.2%; Score 42.5; DB 1; Length 1180;
 Best Local Similarity 57.9%; Pred. No. 53;
 Matches 11; Conservative 1; Mismatches 2; Indels 5; Gaps 1;
 CQ 3 QESTRGXSRG-----RSGR 16
 Db 331 QESTGRNRPNGHPSRSGK 349
 RESULT 14
 FRZ2 DROME STANDARD; PRT; 694 AA.
 AC Q9VAX3; Q94916; Q9VAX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frlizzled protein 2 precursor (Frlizzled-2) (Dfz2).
 GN FZ2 OR CG9739/CG14083.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhanoor P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.;
 RT "A new member of the frizzled family from Drosophila functions as a
 RT WINGLESS receptor.";
 RL Nature 382:225-230(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Peiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolashkov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burrows D., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertey J.M., Cawley S., Dahlke C., Davenport U.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.O., Wei M.-H., Ibegwam C.,
 RA Jaitai B., Kalush F., Karpen G.H., Ke Z., Kennison U.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclib J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard V., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 RL
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins. Inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeleton of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage
 CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC
 CC -1- DOMAIN: Lys-Thr-X-X-Tip motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC
 CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U65569; AAC47273.1; -
 DR EMBL, AE003518; AAF49185.2; -
 DR PIR, S71786; S71786;
 DR FlyBase, FBgn0016797; fz2.
 DR GO, GO:0016021; C:integral to membrane; NMS.
 DR GO, GO:0017147; F:Wnt-protein binding; IDA.
 DR GO, GO:0006928; P:cell motility; IMP.
 DR GO, GO:0007163; P:establishment and/or maintenance of cell po. .; ISS.
 DR GO, GO:0008585; P:female gonad development; IMP.
 DR GO, GO:0016055; P:Wnt receptor signaling pathway; IDA.
 DR InterPro, IPR000539; Frizzled.
 DR InterPro, IPR000832; GPCR secretin.
 DR InterPro, IPR01534; GPCR secretin.
 DR Pfam, PF01393; Frizzled_1.
 DR Pfam, PF01393; Fz_1.
 DR PRINTS, PR00489; FRIZZLED.
 DR SMART, SM00063; FRI1.1.
 DR PROSITE, PS50038; FZ, 1.
 DR PROSITE, PS50261; G-PROTEIN RECD_F2_4; 1.
 KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
 KW Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 694
 FT DOMAIN 23 315
 FT TRANSMEM 316 336
 FT 1 (POTENTIAL).
 FT FRIZZLED PROTEIN 2.
 FT EXTRACELLULAR (POTENTIAL).
 FT 1 (POTENTIAL).

FT DOMAIN 337 352
 FT TRANSMEM 353 373
 FT DOMAIN 374 397
 FT TRANSMEM 398 418
 FT TRANSMEM 419 439
 FT TRANSMEM 440 460
 FT TRANSMEM 461 482
 FT TRANSMEM 483 503
 FT TRANSMEM 504 524
 FT TRANSMEM 535 555
 FT TRANSMEM 556 584
 FT TRANSMEM 585 605
 FT TRANSMEM 606 694
 FT DOMAIN 187 225
 FT DOMAIN 59 180
 FT SITE 608 613
 FT SITE 692 694
 FT CARBOHYD 78 78
 FT CARBOHYD 288 288
 FT CONFLICT 55 55
 FT CONFLICT 417 417
 SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAB096 CRC64;
 Query Match 45.7%; Score 42; DB 1; Length 694;
 Best Local Similarity 52.6%; Pred. No. 35;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHOESTRGSRGSRGSGS 19
 Db 182 SYTEAGSGSGSGSGSGS 200
 RESULT 15
 RN1_FUGRU STANDARD; PRT; 1097 AA.
 ID RN1_FUGRU
 AC Q98TR3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative regulator of nonsense transcripts 1.
 GN RENT1.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 OX NCB1_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clarke D., Elgar G., Clark M.S.;
 RT "Comparative analysis of human 19p12-13 region in Fugu and mouse.";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBSJ databases.
 CC -1- FUNCTION: Eliminates the production of nonsense-containing RNAs
 CC (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the DNA2/NAM7 helicase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AJ301641; CAC33025.1; -
 KW Nonsense-mediated mRNA decay; Hydrolyase; Helicase; ATP-binding;
 KW Zinc-finger.
 FT ZN FING 113 141
 FT ZN FING 165 195
 FT NP BIND 474 481
 FT C2H2-TYPE (ATYPICAL) (POTENTIAL).
 FT C4-TYPE (Potential).
 FT ATP (POTENTIAL).
 FT FDC7C2B4E26AC54 CRC64;
 SQ SEQUENCE 1097 AA; 122338 MW; FDC7C2B4E26AC54 CRC64;

Tue Sep 28 07:11:28 2004

us-09-308-150-2.open.rsp

Page 9

Query Match	45.7%;	Score 42;	DB 1;	Length 1097;
Best Local Similarity	41.7%;	Pred. No. 58;		
Matches 10;	Conservative 3;	Mismatches 5;	Indels 6;	Gaps 1;

```

QY      1 SHQEST-----RGXSRGRSGRSG 18
          |||      ||| :|||
Db      975 SYQGCTNGPAAGRGAMKGKSGRGG 998

```

Search completed: September 28, 2004, 06:05:46
Job time : 8.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-2

Sequence: 1 SHOESTRGSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organellar:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteus:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	797	4	Q16824 homo sapien
2	90	97.8	590	4	Q15206 homo sapien
3	90	97.8	1218	4	Q05331 homo sapien
4	84	91.3	798	4	Q9H4U3 homo sapien
5	84	91.3	1084	4	Q01212 homo sapien
6	75	81.5	591	4	Q01120 homo sapien
7	75	81.5	687	4	Q9H4U2 homo sapien
8	73	79.3	322	4	Q03838 homo sapien
9	70	76.1	465	4	Q03838 homo sapien
10	50	54.3	800	5	Q15845 leishmania
11	49	53.3	213	15	Q7ZB18 chimpanzee
12	48	52.2	327	13	Q09331 callitrichi
13	48	52.2	700	13	Q42378 brachydantio
14	48	52.2	715	13	Q8QF00 brachydantio
15	48	52.2	715	13	Q8QGG8 brachydantio
16	48	52.2	716	13	Q42107 brachydantio

17	48	52.2	822	3	Q9P312 neurospora
18	48	52.2	1229	3	Q86ZM1 neurospora
19	47	51.1	471	5	Q9UAY0 caenorhabdi
20	46	50.0	197	5	Q17626 caenorhabdi
21	46	50.0	203	5	P91026 caenorhabdi
22	46	50.0	236	10	Q9AYR4 oryza sativ
23	46	50.0	236	10	Q7XZ25 oryza sativ
24	46	50.0	353	10	Q8S9W9 oryza sativ
25	46	50.0	399	4	Q9NTA9 oryza sativ
26	46	50.0	467	4	Q9BW18 oryza sativ
27	46	50.0	737	4	Q9B039 oryza sativ
28	46	50.0	836	10	Q9SKZ9 oryza sativ
29	46	50.0	843	12	Q9QNS2 oryza sativ
30	46	50.0	1378	5	Q97405 oryza sativ
31	45	48.9	96	15	Q8JRB5 human immun
32	45	48.9	96	15	Q9WRP4 human immun
33	45	48.9	96	15	Q8JBY4 human immun
34	45	48.9	96	15	Q8ADP5 human immun
35	45	48.9	222	4	Q9E540 human immun
36	45	48.9	243	10	Q9AX45 oryza sativ
37	45	48.9	349	16	Q7U67 rhodospirillum rubrum
38	45	48.9	374	5	Q76561 caenorhabdi
39	45	48.9	476	12	Q80890 herpesvirus
40	45	48.9	652	5	Q77168 apis mellif
41	45	48.9	722	13	Q802M4 brachydantio
42	45	48.9	1096	10	Q94158 oryza sativ
43	45	48.9	1275	4	Q9UQ36 mus muscula
44	45	48.9	1754	11	Q80037 mus muscula
45	45	48.9	1783	4	Q15038 homo sapien

ALIGNMENTS

RESULT 1	Q16824	PRELIMINARY;	PRT;	797 AA.
ID	Q16824			
AC	Q16824			
DT	01-NOV-1996 (TRENBLREL. 01, Created)			
DT	01-NOV-1996 (TRENBLREL. 01, Last sequence update)			
DT	01-UN-2003 (TRENBLREL. 24, Last annotation update)			
DE	Profilaggrin (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91064347; PubMed=2248957;			
RA	Gen S.O.; McBride O.W.; Idler W.W.; Markova N.; Stelner P.M.;			
RT	"Organization, structure, and polymorphisms of the human profilaggrin gene [published erratum appears in Biochemistry 1991 Jun 11;30(23):5814]."			
RL	Biochemistry 29:9432-9440(1990).			
DR	EMBL; M60502; AAA63248.1; -			
DR	GO; GO:0005198; F:structural molecule activity; IEA.			
DR	InterPro; IPR003303; Filaggrin.			
DR	Pfam; PF03515; Filaggrin; 4.			
DR	PRINTS; PR00487; FILAGGRIN.			
FT	NON TER			
SO	SEQUENCE			
Query Match	797 AA;	85176 MW;	60E6184763BDA86B CRC64;	
Best Local Similarity	97.8%;	Score 90;	DB 4;	Length 797;
Matches	18;	Conservative 0;	Mismatches 1;	Indels 0;
QY	1 SHOESTRGSRGRSGSGS 19			
DB	427 SHOESTRGSRGRSGSGS 445			
RESULT 2				

```

Q15206      PRELIMINARY;      PRT;      990 AA.
ID Q15206
AC Q15206;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=9106347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
   gene [published erratum appears in Biochemistry 1991 Jun
   11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
   gene.";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
FT NON_TER      990
SQ SEQUENCE      990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match      97.8%; Score 90; DB 4; Length 990;
Best Local Similarity 94.7%; Pred. No. 2.3e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTRGSRGSRGSRGSGS 19
Db 551 SHOESTRGSRGSRGSRGSGS 569

RESULT 3
ID Q05331      PRELIMINARY;      PRT;      1218 AA.
AC Q05331;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Filaggrin (PROFILAGGRIN) (Fragment).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Markov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
   Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
DE -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
   DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
   TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
DE -1- PFM: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
   HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
   324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
   DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC CC      PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
CC REPEATS.
DR EMBL; M96943; AAA36487.1;
DR PIR; A48118; A48118.
DR HSPB; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_5100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA_BIND      19      32      SITE I (BY SIMILARITY).
FT CA_BIND      62      73      SITE II (BY SIMILARITY).
FT NON_TER      1218
SQ SEQUENCE      1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match      97.8%; Score 90; DB 4; Length 1218;
Best Local Similarity 94.7%; Pred. No. 3.7e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTRGSRGSRGSRGSGS 19
Db 773 SHOESTRGSRGSRGSRGSGS 791

RESULT 4
ID Q9H4U3      PRELIMINARY;      PRT;      798 AA.
AC Q9H4U3;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE D14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lait G.;
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL356504; CAC13171.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; Filaggrin.
FT NON_TER      1
SQ SEQUENCE      798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match      91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 2.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 SHOESTRGSRGSRGSRGSGS 19
Db 428 SHOESTRGSRGSRGSRGSGS 446

RESULT 5
ID Q01212      PRELIMINARY;      PRT;      1084 AA.

```

```

AC 001212; 003840;
AC 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-NOV-2003 (T-EMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
DR NON TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 91.3%; Score 84; DB 4; Length 1084;
Best Local Similarity 89.5%; Pred. No. 3.2e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGRSGS 19
DB 64 SHOESTRGRSGRSGRSGS 82

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR Polymorphism.
FT PROPEP 1 293 POTENTIAL.
FT CHAIN 294 467 FILAGGRIN.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FILAGGRIN.
FT CA BIND 19 32 SITE I (BY SIMILARITY).
FT CA BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.00051;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGRSGS 19
DB 449 SHOESTRGRSGRSGRSGS 467

RESULT 7
ID Q9H4U2 PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE D1J4N1.1.1 (profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; Filaggrin.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 687;
Best Local Similarity 84.2%; Pred. No. 0.0006;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 SHOESTRGXSRGSRGSGS 19
 DB 449 SHOESTRGXSRGSRGSGS 467

RESULT 8

ID 075370 PRELIMINARY; PRT; 322 AA.

AC 075370;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Epidermal filaggrin (Fragment).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;
 RA Garbal-Neuhäuser E., Dirloux U.J., Arnaud M., Dalbon P., Sebbag M.,
 RA Vincent C., Simon M., Senshu T., Maeson-Bessiere C.,
 RA Jolivet-Reynaud C., Jolivet M., Serre G.;
 RT "The epidermal targeted by the rheumatoid arthritis-associated
 RT anti-filaggrin autoantibodies are posttranslationally generated on
 RT various sites of (pro)filaggrin by deamination of arginine residues."
 RL U. Immunol. 162:585-594(1993).
 DR EMBL; AF043380; AAC23559.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 2.
 DR PRINTS; PRO0487; FILAGGRIN.
 FT NON_TER 1
 FT NON_TER 322
 SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230DBFF9E0 CRC64;

Query Match 79.3%; Score 73; DB 4; Length 322;
 Best Local Similarity 83.3%; Pred. No. 0.00057;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRGSGS 18

DB 305 SHOEFTRGSGRSGRSGS 322

RESULT 9

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003838;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE FILAGGRIN (PROFILAGGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 RT gene."
 RL Biochemistry 29:9432-9440(1990).
 RN [2]
 RP REVISIONS
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 RT gene."
 RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
 CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
 CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 CC -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
 CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
 CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin
 CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 DR EMBL; M60499; AAA63246.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 3.
 DR PRINTS; PRO0487; FILAGGRIN.
 FT NON_TER 1
 FT NON_TER 465
 FT NON_TER 465
 SQ SEQUENCE 465 AA; 50280 MW; C683744C5E134097 CRC64;

Query Match 76.1%; Score 70; DB 4; Length 465;
 Best Local Similarity 78.9%; Pred. No. 0.0027;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRGSGS 19

DB 227 SHOESTRGXSRGSRGSGS 245

RESULT 10

ID 015845 PRELIMINARY; PRT; 800 AA.

AC 015845;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE L1439.4.
 GN L1439.4.
 OS Leishmania major.
 CC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_Taxid=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Friledlin;
 RX MEDLINE=99178987; PubMed=10077609;
 RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
 RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
 RA Bastien P., Fu G., Ivans A., Stuart K.;
 RT "Leishmania major Friledlin chromosome 1 has an unusual distribution of
 RT protein-coding genes."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
 DR EMBL; AB001274; AAC24675.1; -;
 DR PIR; T02852; T02852.
 SQ SEQUENCE 800 AA; 83782 MW; DA57AB627913CBAC CRC64;

Query Match 54.3%; Score 50; DB 5; Length 800;
 Best Local Similarity 55.6%; Pred. No. 9.7;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTRGXSRGSRGSGS 19

DB 429 HRDGVRLSTTASGRSGS 446

RESULT 11

ID 072B18 PRELIMINARY; PRT; 213 AA.

AC 072B18;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Vif protein.
 GN VIF.

OS Chimpanzee immunodeficiency virus (SIV(cp2)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxId=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SIMVmd5440;
 RX MEDLINE=22550088; PubMed=12663793;
 RA Hu J., Switzer W.M., Foley B.T., Robertson D.L., Goeken R.N.,
 Korber B.T., Hirsch V.M., Beer B.E.;
 RT "Characterization and Comparison of Recombinant Simian
 Immunodeficiency Virus from Drill (Mandrillus leucophaeus) and
 Mandrill (Mandrillus sphinx) Isolates.";
 RL J. Virol. 77:4867-4880(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SIMVmd5440;
 RA Hu J., Hirsch V.M., Beer B.E.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY159322; AAC22476.1; -;
 DR GO; GO:0019058; P:Viral infectious cycle; IEA.
 DR InterPro; IPR000475; Viral_infect.
 DR Pfam; PF00559; Vif_1
 DR PRINTS; PR00349; VIRIONINPECT.
 DR PRODOM; PD000063; Viral_infect; 1.
 SQ SEQUENCE 213 AA; 24885 MW; EA626C2C0BP4ED27 CRC64;

Query Match 53.3%; Score 49; DB 15; Length 213;
 Best Local Similarity 52.9%; Pred. No. 3 4; Mismatches 5; Indels 0; Gaps 0;
 Matches 9; Conservative 3;

QY 3 QESTRGSRGSRGSGS 19
 Db 163 REETRGARRSRGRGTGA 179

RESULT 12
 Q993H1 PRELIMINARY; PRT; 327 AA.
 ID Q993H1;
 AC Q993H1;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE ORF39.
 OS Callitriche herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxId=106331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT0149;
 RX MEDLINE=21107697; PubMed=11158621;
 RA Wang F., Ramey J., Rivallier P., Quink C., Garber R.L., Beier D.R.,
 Chang F.;
 RT "An Epstein-Barr-related herpesvirus from marmoset lymphomas.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1224-1229(2001).
 DR EMBL; AF319782; AAC38247.1; -;
 DR HSP; P03211; IWH1
 DR GO; GO:0042025; C:host cell nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR004186; EBNA1.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF02905; EBNA1; 1.
 SQ SEQUENCE 327 AA; 34334 MW; 79F03C1D2015C94A CRC64;

Query Match 52.2%; Score 48; DB 12; Length 327;
 Best Local Similarity 75.0%; Pred. No. 8; Mismatches 3; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

QY 7 RGSRGSRGSGS 18
 Db 69 RGSRGSRGSGS 80

RESULT 13
 O42378 PRELIMINARY; PRT; 700 AA.
 ID O42378;
 AC O42378;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE RNA helicase (DEAD box).
 GN VASA OR VLG.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98025484; PubMed=9376327;
 RA Olsen L.C., Aasland R., Fjose A.;
 RT "A vasa-like gene in Zebrafish identifies putative primordial germ
 cells.";
 RL Mech. Dev. 66:95-105(1997).
 CC -1 SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 DR EMBL; Y12007; CAAT2735.1; -;
 DR HSP; Q58083; IHW8.
 DR ZFIN; ZDB-GENE-990415-272; vasa.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000629; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 DR ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 700 AA; 75312 MW; 726B56A90DC2D83 CRC64;

Query Match 52.2%; Score 48; DB 13; Length 700;
 Best Local Similarity 44.4%; Pred. No. 18; Mismatches 6; Indels 8; Gaps 1;
 Matches 12; Conservative 1;

QY 1 SHQSTRGSRGSGS 19
 Db 26 SNSECTEGSWKMTGDSFRGSGRGS 52

RESULT 14
 Q8QF00 PRELIMINARY; PRT; 715 AA.
 ID Q8QF00;
 AC Q8QF00;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Vase-like protein.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vatland A., Olsen L.C.;
 RT "Establishment of transgenic zebrafish lines expressing green
 fluorescent protein in their germ cells.";
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL -1 SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC EMBL; AJ311625; CAC84069.1; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-2

Perfect score: 92
Sequence: 1 SHQESTRXSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29G404:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003ae:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	90	97.8	19	2	AAW61506 Peptide c
2	90	97.8	19	2	AAW61514 Peptide c
3	90	97.8	1467	5	ABH97605 Novel hum
4	86	93.5	19	2	AAW61517 Peptide c
5	85	92.4	19	2	AAW61516 Peptide c
6	84	91.3	19	2	AAW61508 Peptide c
7	84	91.3	19	2	AAW61515 Peptide c
8	84	91.3	19	2	AAW61507 Peptide c
9	84	91.3	19	2	AAW61509 Peptide c
10	84	91.3	19	2	AAW61510 Peptide c
11	84	91.3	19	2	AAW61505 Peptide c
12	78	84.8	19	2	AAW61512 Peptide c
13	78	84.8	19	2	AAW61511 Peptide c
14	78	84.8	19	2	AAW61513 Peptide c
15	75	81.5	21	2	AAW61520 Peptide X
16	75	81.5	22	4	AAW61523 IGP1546 P
17	59	64.1	330	2	AAW22956 Human fil
18	59	64.1	330	2	AAW22955 Human fil
19	59	64.1	330	2	AAW22957 Human fil
20	55	59.8	330	2	AAW22954 Human fil
21	51	55.4	1711	4	AAW79819 Human pro
22	51	55.4	1951	4	AAW78835 Human pro
23	50	54.3	477	6	ABO07142 Novel hum
24	50	54.3	2348	4	ABG10929 Novel hum
25	48	52.2	700	4	AAW02421 Danio rei

26	47	51.1	184	3	AAB57041 Human pro
27	47	51.1	1038	7	ADC03412 Rice flow
28	46	50.0	443	4	AAW40069 Human pol
29	46	50.0	574	4	AAW41855 Human pol
30	46	50.0	752	4	AAW23535 Novel hum
31	46	50.0	19938	6	ABP76682 Streptomy
32	45	48.9	222	4	AAW93652 Human pol
33	45	48.9	1791	6	ABO52986 Human put
34	44	47.8	56	4	AAW60540 Propionib
35	44	47.8	56	6	ABW57059 Propionib
36	44	47.8	134	4	ABG23065 Novel hum
37	44	47.8	157	4	ABG41022 Zea may
38	44	47.8	173	6	ABR41796 Human DIT
39	44	47.8	215	3	AAW46917 Arabidops
40	44	47.8	229	7	ADW59825 Rat prote
41	44	47.8	229	7	ADW59833 Rat prote
42	44	47.8	229	7	ADW59837 Rat prote
43	44	47.8	229	7	ADW59829 Rat prote
44	44	47.8	258	3	AAW46916 Arabidops
45	44	47.8	297	3	AAW46915 Arabidops

ALIGNMENTS

RESULT 1
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX
AC AAW61506;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 9 /note= "Citruiline"
XX
FN W09822503-A2.
XX
PD 28-MAY-1998.
XX
PF 14-NOV-1997; 97WO-NL000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPT, 1998-398613/34.
XX
DR Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
PS Disclosure: Page 6; 19pp; English.
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 1 SHQESTRGXSRGSRGSGS 19

RESULT 2

ID AAW61514 standard; peptide; 19 AA.

XX AAW61514;

XX 26-OCT-1998 (first entry)

XX Peptide cf, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 1.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 1 SHQESTRGXSRGSRGSGS 19

RESULT 3

ABB97605
 ID ABB97605 standard; protein; 1467 AA.

XX ABB97605;

XX 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 873.

KM Human; anti-nausea; vulnary; anti-inflammatory; immunomodulator;
 KW antineurotic; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 XX expressed sequence tag.

XX Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-292408/33.

XX N-PSDB; ABB32791.

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

XX Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.4e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHQESTRGXSRGSRGSGS 19
 |||||
 DB 773 SHQESTRGXSRGSRGSGS 791

RESULT 4

ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

XX 26-OCT-1998 (first entry)

XX Peptide cfo, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX

```

OS Synthetic.
OS Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97MO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 93.5%; Score 86; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 7.8e-07;
XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 SHQESTRGXSRGRSGSGS 19
XX |||||:|||||
XX 1 SHQESTGRGRSGRSGSGS 19
XX
XX Db
XX
XX RESULT 5
XX AAW61516
XX ID AAW61516 standard; peptide; 19 AA.
XX
XX AC AAW61516;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE Peptide cfr, based on cDNA of a profilaggrin repeat.
XX
XX KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX PN WO9822503-A2.
XX
XX PD 28-MAY-1998.
XX
XX PF 14-NOV-1997; 97MO-NL000624.
XX
XX PR 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX

```

```

XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 92.4%; Score 85; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.1e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 SHQESTRGXSRGRSGSGS 19
XX |||||:|||||
XX 1 SHQESTGRGRSGRSGSGS 19
XX
XX Db
XX
XX RESULT 6
XX AAW61508
XX ID AAW61508 standard; peptide; 19 AA.
XX
XX AC AAW61508;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE Peptide cfr4, based on cDNA of a profilaggrin repeat.
XX
XX KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 13
XX FT /note= "Citru1line"
XX
XX PN WO9822503-A2.
XX
XX PD 28-MAY-1998.
XX
XX PF 14-NOV-1997; 97MO-NL000624.
XX
XX PR 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX

```

CC Patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTRGXRGSRGRSGS 19
 |||||
 DB 1 SHQESTRGXRGSRGRSGS 19

RESULT 7
 AAW61515
 ID AAW61515 standard; peptide; 19 AA.

AC AAW61515;

DT 26-OCT-1998 (first entry)

DE Peptide cFA, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX

XX

DB 1 SHQESTAGSRGRGRSGS 19

RESULT 8
 AAW61507
 ID AAW61507 standard; peptide; 19 AA.

AC AAW61507;

DT 26-OCT-1998 (first entry)

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.

XX Synthetic.
OS Homo sapiens.

XX Key Location/Qualifiers
FH Modified-site 16
FT /note= "Citruilline"
XX

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

XX Query Match 91.3%; Score 84; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
DB 1 SHQESTRGXSRGRSGSGS 19

RESULT 10
AAW61510
ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

XX 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.

XX Synthetic.
OS Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 7

FT /note= "Citruilline"
FT Modified-site 9
FT /note= "Citruilline"
XX

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

XX Query Match 91.3%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 1.6e-06;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGXSRGRSGSGS 19
DB 1 SHQESTRGXSRGRSGSGS 19

RESULT 11
AAW61505
ID AAW61505 standard; peptide; 19 AA.

XX AAW61505;

XX 26-OCT-1998 (first entry)

XX Peptide cfc1, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.

XX Synthetic.
OS Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 7

XX /note= "Citruilline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX Sequence 19 AA;
 SQ
 Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SHQSTKGRSGRSGSGS 19
 DB 1 SHQSTKGRSGRSGSGS 19
 RESULT 12
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 XX AAW61512;
 AC
 XX 26-OCT-1998 (first entry)
 DT
 XX
 DE Peptide cfc8, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citrulline"
 FT Modified-site 13 /note= "Citrulline"
 FT
 XX W09822503-A2.
 EN
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX Sequence 19 AA;
 SQ
 Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 SHQSTKGRSGRSGSGS 19
 DB 1 SHQSTKGRSGRSGSGS 19
 RESULT 13
 AAW61511
 ID AAW61511 standard; peptide; 19 AA.
 XX
 XX AAW61511;
 AC
 XX 26-OCT-1998 (first entry)
 DT
 XX
 DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citrulline"
 FT Modified-site 11 /note= "Citrulline"
 FT
 XX W09822503-A2.
 EN
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGRSGSGS 19
 Db 1 SHQESTXGRSGRSGSGS 19

RESULT 14
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.

AC AAW61513;
 DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruilline"
 FT Modified-site 16
 FT /note= "Citruilline"

PN W09822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.5e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGRSGSGS 19
 Db 1 SHQESTXGRSGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 DT 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 9
 FT /note= "Citruilline"

PN W09822503-A2.
 PD 28-MAY-1998.
 PF 14-NOV-1997; 97WO-NL000624.
 PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 4.9e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HQESTRGSRGRSGSGS 19
 Db 4 HQESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:11
 Job time : 56.1302 secs

Tue Sep 28 07:33:33 2004

us-09-308-150-2.open.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-2
Perfect score: 92
Sequence: 1 SHOESTRGSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-2
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-3
7	84	91.3	19	9	US-09-308-150-4
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-6
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-7
12	78	84.8	19	9	US-09-308-150-8
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	50	54.3	477	15	US-10-161-927-62	Sequence 62, Appl
17	48	52.2	506	16	US-10-437-963-185974	Sequence 185974,
18	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
19	47	51.1	257	16	US-10-437-963-105967	Sequence 105967,
20	46.5	50.5	176	16	US-10-437-963-196023	Sequence 196023,
21	46	50.0	16	10	US-09-791-524-150	Sequence 150, App
22	46	50.0	130	16	US-10-437-963-104588	Sequence 104588,
23	46	50.0	224	12	US-10-424-599-269475	Sequence 269475,
24	46	50.0	226	16	US-10-437-963-102742	Sequence 102742,
25	46	50.0	445	12	US-10-424-599-267922	Sequence 267922,
26	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
27	46	50.0	1087	16	US-10-641-991-24	Sequence 8, Appl
28	46	50.0	19608	15	US-10-084-846A-8	Sequence 30865, A
29	45	48.9	79	14	US-10-029-366-103885	Sequence 103033,
30	45	48.9	213	16	US-10-437-963-103033	Sequence 1630387,
31	45	48.9	294	16	US-10-437-963-169387	Sequence 169387,
32	45	48.9	563	16	US-10-437-963-198539	Sequence 198539,
33	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
34	45	48.9	2263	16	US-10-408-765A-2231	Sequence 2231, Ap
35	44.5	48.4	307	12	US-10-425-114-39059	Sequence 39059, A
36	44	47.8	78	16	US-10-437-963-203073	Sequence 203073,
37	44	47.8	430	16	US-10-437-963-164039	Sequence 164039,
38	44	47.8	511	16	US-10-437-963-188040	Sequence 188040,
39	44	47.8	1071	16	US-10-437-963-118389	Sequence 118389,
40	43.5	47.3	85	12	US-10-424-599-161928	Sequence 161928,
41	43.5	47.3	181	16	US-10-437-963-117640	Sequence 117640,
42	43.5	47.3	209	16	US-10-437-963-175682	Sequence 175682,
43	43	46.7	50	14	US-10-029-386-27692	Sequence 27692, A
44	43	46.7	76	12	US-10-424-599-248912	Sequence 248912,
45	43	46.7	87	9	US-09-867-550-764	Sequence 764, App

ALIGNMENTS

RESULT 1
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Walthers Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozet Maria Hendrik
APPLICANT: Hoel, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.86-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 2

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.8e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 3

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 19

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
 ; CURRENT APPLICATION NUMBER: US/09/308.150
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL 1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGRSGSGS 19
 DB 1 SHOESTRGXSRGRSGSGS 19

RESULT 5

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-1

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 6
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-3

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 7
US-09-308-150-4

Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-4

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGXSRGSRSGS 19
DB 1 SHOESTRGXSRGSRSGS 19

RESULT 8
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-5

US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQSTXGXSRRGSGSGS 19
Db 1 SHQSTXGXSRRGSGSGS 19

RESULT 9

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.1e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQSTXGXSRRGSGSGS 19
Db 1 SHQSTXGXSRRGSGSGS 19

RESULT 10

US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQSTXGXSRRGSGSGS 19
Db 1 SHQSTXGXSRRGSGSGS 19

RESULT 11

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHQSTXGXSRRGSGSGS 19
Db 1 SHQSTXGXSRRGSGSGS 19

RESULT 12

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGRSGRSGS 19
Db 1 SHOESTRGSRGRSGRSGS 19

```

```

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcoij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGRSGRSGS 19
Db 1 SHOESTRGSRGRSGRSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcoij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00045;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 HOESTXGSRGRSGRSGS 19
Db 4 HOESTXGSRGRSGRSGS 21

```

```

RESULT 15
US-09-747-0294-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NPU500 INNS:031
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US/09/747,029A
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```


Tue Sep 28 07:11:27 2004

us-09-308-150-2.open.ral

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)

67.892 Million cell updates/sec

Title: US-09-308-150-2

Perfect score: 92

Sequence: 1 SHOESTRGXSGRSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	51.1	406	4	US-09-252-991A-1857
2	47	51.1	421	4	US-09-252-991A-32326
3	46	50.0	465	4	US-09-252-991A-18919
4	45	48.9	204	4	US-09-252-991A-21317
5	44	47.8	369	4	US-09-252-991A-20245
6	44	47.8	483	4	US-09-252-991A-19015
7	44	47.8	566	4	US-09-252-991A-24514
8	43	46.7	201	4	US-09-252-991A-25745
9	43	46.7	508	4	US-09-252-991A-28613
10	43	46.7	653	4	US-09-513-057C-13
11	43	46.7	730	4	US-08-591-502B-49
12	43	46.7	842	4	US-08-591-502B-50
13	43	46.7	842	4	US-08-591-502B-51
14	43	46.7	843	4	US-09-719-528A-2
15	43	46.7	843	4	US-08-591-502B-45
16	43	46.7	843	4	US-08-591-502B-48
17	43	46.7	843	4	US-08-591-502B-59
18	43	46.7	845	4	US-08-591-502B-46
19	42	45.7	138	4	US-09-252-991A-16961
20	42	45.7	169	3	US-09-342-084-6
21	42	45.7	250	4	US-09-252-991A-26599
22	42	45.7	274	4	US-09-252-991A-16696
23	42	45.7	629	4	US-09-252-991A-22901
24	42	45.7	703	4	US-09-252-991A-20353
25	42	45.7	2237	3	US-08-354-973-1
26	41	44.6	143	3	US-08-868-699A-4
27	41	44.6	143	4	US-09-757-014-4

28	41	44.6	205	4	US-09-252-991A-18529	Sequence 18529, A
29	41	44.6	219	4	US-09-621-976-4221	Sequence 4221, Ap
30	41	44.6	255	4	US-09-252-991A-19598	Sequence 19598, A
31	41	44.6	286	4	US-09-252-991A-22717	Sequence 22717, A
32	41	44.6	354	4	US-09-252-991A-29717	Sequence 29717, A
33	41	44.6	404	4	US-09-489-039A-11606	Sequence 11606, A
34	41	44.6	409	4	US-09-489-039A-14217	Sequence 14217, A
35	41	44.6	446	4	US-09-252-991A-21791	Sequence 21791, A
36	41	44.6	539	4	US-09-252-991A-17070	Sequence 17070, A
37	41	44.6	575	3	US-08-922-865-2	Sequence 2, Appli
38	41	44.6	575	4	US-09-510-949-2	Sequence 2, Appli
39	41	44.6	722	4	US-09-984-890-4	Sequence 4, Appli
40	41	44.6	724	4	US-09-984-890-2	Sequence 2, Appli
41	41	44.6	745	4	US-09-523-849-36	Sequence 36, Appli
42	41	44.6	796	3	US-08-868-699A-2	Sequence 2, Appli
43	41	44.6	796	4	US-09-757-014-2	Sequence 2, Appli
44	41	44.6	803	4	US-09-252-991A-23614	Sequence 23614, A
45	40.5	44.0	255	3	US-09-342-084-11	Sequence 11, Appli

ALIGNMENTS

```
RESULT 1
US-09-252-991A-19857
; Sequence 19857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/034,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19857
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19857

Query Match      51.1%, Score 47; DB 4; Length 406;
Best Local Similarity 50.0%; Pred. No. 8.7;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      2 HOESTRGXSGRSGSGS 19
DB      127 HGQSRGDRAGRGHGA 144

RESULT 2
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
```


Tue Sep 28 07:11:27 2004

us-09-308-150-2.open.ra1

Page 3

;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 24514
;; LENGTH: 586
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24514

Query Match 47.8%; Score 44; DB 4; Length 586;
Best Local Similarity 61.1%; Pred. No. 37;
Matches 11; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 3 OESTR-GXSRGSGSG 18
Db 245 QRSRRTGGGRSGSG 262

RESULT 8
US-09-252-991A-25745
; Sequence 25745, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25745
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25745

Query Match 46.7%; Score 43; DB 4; Length 201;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 HOESTRGXSRGSGR 16
Db 111 HRSRWGAGRGGR 125

RESULT 9
US-09-252-991A-28613
; Sequence 28613, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28613
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28613

Query Match 46.7%; Score 43; DB 4; Length 508;
Best Local Similarity 44.4%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHOESTRGXSRGSGR 18
Db 106 ARRSRRGAGSRGTRG 123

RESULT 10
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 46.7%; Score 43; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 59;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SHOESTRGXSRGSGR 19
Db 573 SRQVSTASASGRGIGS 591

RESULT 11
US-08-591-502B-49
; Sequence 49, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Laver

REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-591-502B-49

Query Match 46.7%; Score 43; DB 4; Length 730;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGSGRSGS 242

RESULT 12
US-08-591-502B-50
Sequence 50, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:

US-08-591-502B-50

Query Match 46.7%; Score 43; DB 4; Length 842;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGSGRSGS 242

RESULT 13
US-08-591-502B-51
Sequence 51, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51

Query Match 46.7%; Score 43; DB 4; Length 842;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGSGRSGS 19
:|||||
Db 233 ARGSGRSGS 242

RESULT 14
US-09-719-528A-2

Sequence 2, Application US/09719528A
Patent No. 6558675
GENERAL INFORMATION:
APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi
Chen, Wei Ning
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
US: THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford U.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2

Query Match 46.7%; Score 43; DB 4; Length 843;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGRSGRSGS 19
:|||||
Db 233 ARKSGRSGS 242

RESULT 15
US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45

Query Match 46.7%; Score 43; DB 4; Length 843;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 SRGRSGRSGS 19
:|||||
Db 233 ARKSGRSGS 242

Search completed: September 28, 2004, 06:26:41
Job time : 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHOESTRGSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	416	2 A32947	filaggrin precursor
2	91	98.9	2248	2 A35938	profilaggrin - hum
3	82	89.1	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2ML47	E2 protein - human
5	54	58.7	471	2 T33997	hypothetical prote
6	50	54.3	123	2 T16234	hypothetical prote
7	50	54.3	800	2 T02852	probable membrane
8	48	52.2	822	2 T51049	related to nucleol
9	46	50.0	197	2 T18918	hypothetical prote
10	46	50.0	203	2 C87801	protein C10G11.9 l
11	46	50.0	306	2 T21220	hypothetical prote
12	46	50.0	373	2 T02976	probable DNA bindi
13	46	50.0	399	2 T46259	hypothetical prote
14	46	50.0	836	2 G84727	probable DNA topol
15	46	50.0	1105	2 T47582	hypothetical prote
16	45	48.9	374	2 T33328	hypothetical prote
17	45	48.9	1337	2 T30291	dextranase - strep
18	44	47.8	89	2 H87367	hypothetical prote
19	44	47.8	229	2 JC7219	nuclear protein SR
20	44	47.8	867	2 T27136	hypothetical prote
21	44	47.8	871	2 T27135	hypothetical prote
22	44	47.8	1496	2 T00499	probable retrolelm
23	43.5	47.3	463	2 T51194	hypothetical prote
24	43	46.7	70	4 T52626	hypothetical FUS/E
25	43	46.7	286	4 S50855	neurotrophin-6 - s
26	43	46.7	462	2 T15543	hypothetical prote
27	43	46.7	462	2 S33798	FUS/CHOP mutant fu
28	43	46.7	526	1 S33799	RNA-binding protei
29	43	46.7	528	2 G02127	fus-like protein -

30	43	46.7	1232	2 S40766	hypothetical prote
31	43	46.7	1415	1 EDBEGA	immediate-early pr
32	43	46.7	1507	2 B47328	natural killer cel
33	42	45.7	150	2 T84777	very hypothetical
34	42	45.7	203	2 T25916	hypothetical prote
35	42	45.7	649	2 C71283	probable ATP-depen
36	42	45.7	694	2 S71786	wingless receptor
37	42	45.7	849	2 A96592	hypothetical prote
38	42	45.7	1015	2 T42013	frequency clock pr
39	42	45.7	1287	2 S55954	viral mRNA transla
40	42	45.7	2233	2 S63347	acetyl-CoA carboxy
41	41.5	45.1	79	2 A40459	nuclear phosphopro
42	41	44.6	350	2 A44843	EMAP30-47 antigen
43	41	44.6	136	2 T35632	probable transpos
44	41	44.6	250	2 T29344	hypothetical prote
45	41	44.6	311	2 T45683	hypothetical prote

ALIGNMENTS

```

RESULT 1
A32947
filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of cl
A/Reference number: A32947; MUID:89296901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-references: GB:M24355; NID:G182604; PID:AAA52454.1; PID:G182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unaassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: Ep hand; epidermis; polymorphism; tandem repeat

Query Match          98.9%; Score 91; DB 2; Length 416;
Best Local Similarity 94.7%; Pred. No. 1.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHOESTRGSRGSRGSGS 19
DB 7 SHOESTRGSRGSRGSGS 25
|||||
|||

RESULT 2
A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; MUID:91064347; PMID:2248957
A/Status: preliminary; not compared with conceptual translation
A/Accession: A35938
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:J02929
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unaassigned calmodulin-related proteins; calmodulin repeat homology

```

C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:246-569/Region: filaggrin repeat
 F:570-893/Region: filaggrin repeat
 F:1074-1397/Region: filaggrin repeat
 F:1573-1896/Region: filaggrin repeat

Query Match 98.9%; Score 91; DB 2; Length 2248;
 Best Local Similarity 94.7%; Pred. No. 7.3e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 1055 SHOESTRGRSGRSGSGS 1073

RESULT 3

A:Accession: A45135
 A:Protein: filaggrin - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C/Accession: A45135
 R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nirmalskirti, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992
 A>Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717
 A:Accession: A45135
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101069; GB:M90967; NID:g190408; PIDN:AAA60177.1; PID:g553621
 A>Note: sequence extracted from NCBI backbone (NCBIP:118773)
 C/Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-q22
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 89.1%; Score 82; DB 2; Length 591;
 Best Local Similarity 89.5%; Pred. No. 5.1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 449 SHOESTRGRSGRSGSGS 467

RESULT 4

W2ML47
 E2 protein - human papillomavirus type 47
 C/Species: human papillomavirus type 47
 A>Note: host Homo sapiens (man)
 C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C/Accession: D35324
 R:Kiyo, T.; Adachi, A.; Ishibashi, M.
 Virology 177, 401-405, 1990
 A>Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
 A:Reference number: A55324; MUID:90281611; PMID:2162112
 A:Accession: D35324
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M2305; NID:g333062; PIDN:AAA46979.1; PID:g333067
 C:Superfamily: papillomavirus E2 protein
 C/Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.22;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 1055 SHOESTRGRSGRSGSGS 1073

DB 342 SRENTGRGRGRGRGSGS 360

RESULT 5

T33997
 hypothetical protein W03G1.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C/Accession: T33997
 R:Pauley, A.; Scheet, P.; Harper, M.
 submitted to the EMBL Data Library, February 1999
 A>Description: The sequence of C. elegans cosmid W03G1.
 A:Reference number: 221454
 A:Accession: T33997
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-471 <PAU>
 A:Cross-references: EMBL:AF125964; PIDN:AAAD14753.1; GSPDB:GN00022; CESP:W03G1.5
 A:Experimental source: strain Bristol N2; clone W03G1
 C/Genetics:
 A:Gene: CESP:W03G1.5
 A:Map position: 4

Query Match 58.7%; Score 54; DB 2; Length 471;
 Best Local Similarity 57.9%; Pred. No. 0.84;
 Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 |||||
 DB 167 SRSRGRGRGRGRGSGS 185

RESULT 6

T16234
 hypothetical protein F32A5.7 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C/Accession: T16234
 R:Pauley, A.
 submitted to the EMBL Data Library, July 1995
 A>Description: The sequence of C. elegans cosmid F32A5.
 A:Reference number: 218482
 A:Accession: T16234
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-123 <PAU>
 A:Cross-references: EMBL:U20864; NID:g669026; PIDN:AA046661.1; CESP:F32A5.7
 A:Experimental source: strain Bristol N2
 C/Genetics:
 A:Gene: CESP:F32A5.7
 A:Introns: 2/3; 49/3; 75/3

Query Match 54.3%; Score 50; DB 2; Length 123;
 Best Local Similarity 56.2%; Pred. No. 0.98;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTRGRSGRSGRSGS 18
 |||||
 DB 93 RQSRGRSGRSGRSGS 108

RESULT 7

T02852
 probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)
 C/Species: Leishmania major
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C/Accession: T02852; H81462
 R:Myler, P.J.
 submitted to the EMBL Data Library, May 1998
 A>Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
 A:Reference number: Z14740
 A:Accession: T02852
 A>Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
 A/Residues: 1-800 <MTL>
 A/Cross-references: EMBL:AB001274; NID:G3264850; PID:G2266920
 R/Wyler, P.J.; Audleman, T.J.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A/Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A/Reference number: AB1455; MIMD:9918987; PMID:10077609
 A/Accession: H81462
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-800 <PIL>
 A/Cross-references: GB:AE001274; NID:G3264850; PIDN:AA24675.1; PID:G2266920; GSPDB:GN00
 A/Experimental source: strain MHOM/IL/81/Friedlin
 C/Genetics:
 A/Gene: L1439.4
 A/Map position: 1

Query Match 54.3%; Score 50; DB 2; Length 800;
 Best Local Similarity 55.6%; Pred. No. 5.7;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTRGSRXGRSGSGS 19
 ||:|||||
 Db 429 HRDVGRLSTTASGRSGS 446

RESULT 8
 T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N/Alternate names: protein B12P1.10
 C/Species: Neurospora crassa
 C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 R/Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A/Accession: T51049
 A/Reference number: Z25286
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-822 <SCH>
 A/Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12P1.10
 A/Experimental source: BAC clone B12P1; strain CR74A
 C/Genetics:
 A/Gene: NCSP:B12P1.10
 A/Map position: 6
 A/Introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 12;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGSRXGRSGSGS 19
 ||:|||||
 Db 414 RESASGRTRGRGRGRT 430

RESULT 9
 T18918
 hypothetical protein C04G2.8 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T18918
 R/Hembry, C.
 submitted to the EMBL Data Library, April 1996
 A/Reference number: Z19045
 A/Accession: T18918
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-197 <MTL>
 A/Cross-references: EMBL:Z70718; PIDN:CAA94670.1; GSPDB:GN00022; CESP:C04G2.8
 A/Experimental source: clone C04G2
 C/Genetics:
 A/Gene: CESP:C04G2.8

A/Map position: 4
 A/Introns: 70/2

Query Match 50.0%; Score 46; DB 2; Length 197;
 Best Local Similarity 52.6%; Pred. No. 6.3;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 ||:|||||
 Db 104 SRRSSRGSVYGRTRGRS 122

RESULT 10
 C87801
 protein C10G11.9 [imported] - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C/Accession: C87801
 R/anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A/Reference number: A75000; MIMD:99069613; PMID:9851918
 A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg
 A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A/Accession: C87801
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-203 <STO>
 A/Cross-references: GS:chr_1; PIDN:ABA42250.1; PID:G1825626; GSPDB:GN00019; CESP:C10G11.5
 A/Genetics:
 A/Gene: C10G11.9
 A/Map position: 1

Query Match 50.0%; Score 46; DB 2; Length 203;
 Best Local Similarity 52.6%; Pred. No. 6.5;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 ||:|||||
 Db 110 SRRSSRGSVYGRTRGRS 128

RESULT 11
 T21220
 hypothetical protein F21H7.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C/Accession: T21220
 R/Gardner, A.
 submitted to the EMBL Data Library, March 1997
 A/Reference number: Z19393
 A/Accession: T21220
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-306 <MTL>
 A/Cross-references: EMBL:Z93379; PIDN:CAB07568.1; GSPDB:GN00023; CESP:F21H7.5
 A/Experimental source: clone F21H7
 C/Genetics:
 A/Gene: CESP:F21H7.5
 A/Map position: 5
 A/Introns: 28/3; 146/3; 232/3
 C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 50.0%; Score 46; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 9.5;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSG 18
 ||:|||||
 Db 48 SHRSKTKMKSGKSGKSG 65

RESULT 12

T02976
 Probable DNA binding protein PCF2 - rice
 C:Species: Oryza sativa (rice)
 C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C:Accession: T02976
 R:Kosugi, S.; Ohashi, Y.
 Plant Cell 9, 1607-1619, 1997
 A:Title: PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell
 A:Reference number: 214803; PMID:97480096; PMID:9338963
 A:Accession: T02976
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-373 <KOS>
 A:Cross-references: EMBL:D87261; NID:G2580439; PID:BAA23143.1; PID:G2580440
 A:Experimental source: cultivar Nipponbare

Query Match 50.0%; Score 46; DB 2; Length 373;
 Best Local Similarity 53.3%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSG 15
 DB 334 SHEQRRGRKRGNG 348

RESULT 13

hypothetical protein DKFZP761E0323.1 - human (fragment)
 C:Species: Homo sapiens (mat)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
 C:Accession: T46259
 R:Blöcker, H.; Beecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223032
 A:Accession: T46259
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-399 <AAA>
 A:Cross-references: EMBL:AL137423
 A:Experimental source: adult amygdala; clone DKFZP761E0323
 C:Genetics:
 A>Note: DKFZP761E0323.1

Query Match 50.0%; Score 46; DB 2; Length 399;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSG 18
 DB 336 SNSRQSGWSGSGRSG 353

RESULT 14

probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84727
 R:Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKien, S.E.; Umayam, L.; Tallon, L.;
 Euss, D.; Nieman, W.C.; White, C.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-769, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: G84727
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-836 <STO>
 A:Cross-references: GB:AE02093; NID:G4263718; PID:AMD15404.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AC2G32000
 A:Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
 Best Local Similarity 69.2%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 7 RGRSGRSGRSG 19
 DB 812 RGRSGRSGRSG 824

RESULT 15

hypothetical protein F24B22.190 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47582
 R:Blöcker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quettler, F.; Salanoubat, M.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 223016
 A:Accession: T47582
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1105 <BLO>
 A:Cross-references: EMBL:AL132957
 A:Experimental source: cultivar Columbia; BAC clone F24B22
 C:Genetics:
 A:Map position: 3
 A:introns: 35/3; 56/2; 294/3; 318/3; 349/3; 376/2; 426/3; 455/1; 485/3; 508/3; 568/3; 633/3
 A>Note: F24B22.190

Query Match 50.0%; Score 46; DB 2; Length 1105;
 Best Local Similarity 57.9%; Pred. No. 32;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

CY 1 SHQSTRGRSGRSGRSG 19
 DB 250 SRSRSGRSGRSGRSG 268

Search completed: September 28, 2004, 06:14:58
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-3

Perfect score: 92
Sequence: 1 SHQESTGRGXGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	91	98.9	1	FILA_HUMAN
2	58	63.0	2	P20930 homo sapien
3	50	54.3	3	P22420 human papil
4	45	48.9	4	LSM4_CABEL
5	43	46.7	5	DEXT_STRDO
6	43	46.7	6	EBP2_CABEL
7	43	46.7	7	FUS_BOVIN
8	43	46.7	8	FUS_MOUSE
9	43	46.7	9	FUS_HUMAN
10	43	46.7	10	YOO5_CABEL
11	43	46.7	11	ICP4_HSVMG
12	43	45.7	12	NCRC_MOUSE
13	42	45.7	13	YDPC_SCHPO
14	42	45.7	14	FRZ2_DROME
15	42	45.7	15	FRQ_CRESP
16	42	45.7	16	SK12_YEAST
17	42	45.7	17	COAC_YEAST
18	41	44.6	18	CMC1_EIMAC
19	41	44.6	19	RIX21_DROME
20	41	44.6	20	2ABA_YEAST
21	41	44.6	21	WR11_CABEL
22	41	44.6	22	SE5_RAT
23	41	44.6	23	SON_MOUSE
24	41	44.6	24	2426 1
25	40.5	44.0	25	RNPH_MYCTU
26	40.5	44.0	26	AN3_XENLA
27	40.5	44.0	27	FBR1_MOUSE
28	40.5	44.0	28	AP50_DICDI
29	40.5	44.0	29	WRK2_MOUSE
30	39.5	42.9	30	SPS_SOLTU
31	39.5	42.9	31	SFR7_HUMAN
32	39.5	42.9	32	L2M1_ADB40
33	39.5	42.9	33	DH1_MAZE
34	39.5	42.9	34	GRP1_ARTSA

ALIGNMENTS

34	39	42.4	315	1	SOL2_YEAST	P37262 saccharomyc
35	39	42.4	325	1	FRZB_BOVIN	O95117 bos taurus
36	39	42.4	484	1	CLK1_HUMAN	P49759 homo sapien
37	39	42.4	562	1	MTRB_MYCLE	O9ccj1 mycobacteri
38	39	42.4	638	1	R220_HUMAN	O01546 homo sapien
39	39	42.4	983	1	Y144_HUMAN	Q14157 homo sapien
40	39	42.4	1130	1	SN3B_HUMAN	O75182 homo sapien
41	39	42.4	1527	1	ARHB_RAT	O96667 rattus norv
42	38.5	41.8	135	1	PRH1_DROME	O02427 drosophila
43	38.5	41.8	223	1	PRH1_MOUSE	P37439 mus musculu
44	38.5	41.8	259	1	RNPH_MYCLE	P37939 mycobacteri
45	38.5	41.8	375	1	SR55_DROME	P26686 drosophila

RESULT 1
FILA_HUMAN
ID FILA_HUMAN STANDARD: PRT; 416 AA.
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE=89296901; PubMed=2740331;
RA Wekinley-Grant L.C., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinhart P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RN [2]
RP CITRULINATION.
RX MEDLINE=96374389; PubMed=8780679;
RA Senhu T., Kan S., Ogawa H., Manabe M., Asaga H.;
RT "Preferential deimination of keratin K1 and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -I- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -I- PTM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -I- PTM: Undergoes deimination of some arginine residues
CC (citrullination).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL laboratory
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24355; AA52454.1; -
CC
CC Genew; HGNC:3748; FLG.
CC MIM; 135940; -
CC DR GO; GO:0005882; C:intermediate filament; NAS.
CC DR GO; GO:0005198; P:structural molecule activity; NAS.
CC DR GO; GO:0007275; P:development; NAS.
CC DR InterPro; IPR003303; Filaggrin.

DR Pfam: PF03516; Flaggrin; 2.
 DR PRINTS: P000487; FLAGGRIN.
 KM Phosphorylation; Citrullination; Developmental protein.
 FT NON-TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 98.9%; Score 91; DB 1; Length 416;
 Best Local Similarity 94.7%; Pred. No. 4e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGSGS 19
 DB 7 SHQSTRGRSGRSGSGS 25

RESULT 2
 VE2_HP47 STANDARD; PRT; 506 AA.
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 RT type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).
 CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCCNNNNAGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M32305; AAA46979.1; -.
 DR PIR: D35324; W2WL47.
 DR HSSP: P03122; 2BOP.
 DR InterPro: IPR000427; E2_C.
 DR InterPro: IPR001866; E2_N.
 DR InterPro: IPR009021; Viral_DNA_bd.
 DR Pfam: PF00511; E2_C; 1.
 DR Pfam: PF00508; E2_N; 1.
 DR ProDom: PD000672; E2_C; 1.
 DR ProDom: PD000678; E2_N; 1.
 KM Early protein; Transcription regulation; Activator; DNA-binding;
 KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF75065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.077;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGSGS 19

DB 342 SREGNTRGRGRGRGRAGS 360

RESULT 3
 LSM4_CABEL STANDARD; PRT; 123 AA.
 ID LSM4_CABEL
 AC Q19952;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable U6 snRNA-associated Sm-like protein LSM4.
 GN F32A5.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Rhabditidae; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
 CC snRNA (By similarity).
 CC -1- SUBUNIT: LSM subunits form a heteromer with a doughnut shape (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U0864; AAC46661.1; -.
 DR PIR: T16234; T16234.
 DR WormPeP: F32A5.7; CE01277.
 DR InterPro: IPR006649; snRNP.
 DR InterPro: IPR001163; snRNP_Sm.
 DR Pfam: PF01423; LSM; 1.
 DR ProDom: PD020287; snRNP; 1.
 DR SMART: SM00651; Sm; 1.
 KM Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
 KM RNA-binding.
 SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 54.3%; Score 50; DB 1; Length 123;
 Best Local Similarity 56.2%; Pred. No. 0.3;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGRSGRSGSGS 18
 DB 93 REOSRGRGGRGGRGG 108

RESULT 4
 DEXT_STRDO STANDARD; PRT; 1337 AA.
 ID DEXT_STRDO
 AC P39653;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
 DE glucanohydrolase).
 GN DEX.
 OS Streptococcus downei (Streptococcus sobrinus).
 OG Plaamid PYA902.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1317;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
 RC STRAIN=6715 / UAB66;
 RX MEDLINE=94292401; PubMed=8021165;
 RA Wanda S.-Y., Curtiss R. III;
 RT "Purification and characterization of Streptococcus sobrinus
 RT dextranase produced in recombinant Escherichia coli and sequence
 RT analysis of the dextranase gene";
 RL J. Bacteriol. 176:3839-3850(1994).
 CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
 CC PELLICULE-COATED TOOTH SURFACE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
 CC linkages in dextran.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (potential).
 CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
 CC and at 39 degrees Celsius.
 CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M96978; AAA2172.1; -;
 CC InterPro; IPR001899; Gram_pos_anchor.
 CC Pfam; PF00746; Gram_pos_anchor; 1.
 CC TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 CC PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
 CC HydroLase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
 CC Signal; PfamId. 1
 CC FT SIGNAL 30
 CC FT CHAIN 1 1308 DEXTRANASE.
 CC FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
 CC FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).
 CC FT MOD RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 CC SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A223D0 CRC64;
 CC
 CC Query Match 48.9%; Score 45; DB 1; Length 1337;
 CC Best Local Similarity 47.1%; Pred. No. 24;
 CC Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 1 SHOESTGRXGRGRG 17
 CC Db 1154 ANDSTKSSADSGKS 1170
 CC
 CC RESULT 5
 CC EBP2_CABEL
 CC ID EBP2_CABEL STANDARD; PRT; 340 AA.
 CC AC Q09958;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Problematic RNA processing protein EBP2 homolog.
 CC GN C18A3.3.
 CC OS Caenorhabditis elegans.
 CC CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC CC Rhabditidae; Peleodermidae; Caenorhabditis.
 CC CC NCB1_TaxID=6239;
 CC RN NCB1 [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Bristol N2;
 CC RA Hallsworth K.;
 CC RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC RN [2]
 CC RP REVISIONS.
 CC RA Waterston R.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Required for the processing of the 27S pre-rRNA (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -1- SIMILARITY: Belongs to the EBP2 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U28944; AA68370.2; -;
 CC WormRep; C18A3.3; CE28850.
 CC InterPro; IPR008610; EBP2.
 CC Pfam; PF05890; EBP2; 1.
 CC DR Hypothetical protein; Ribosome biogenesis; Nuclear protein;
 CC KW Coiled coil.
 CC FT DOMAIN 206 245 COILED COIL (POTENTIAL).
 CC SQ SEQUENCE 340 AA; 38083 MW; 13BE25D985A5D55 CRC64;
 CC
 CC Query Match 46.7%; Score 43; DB 1; Length 340;
 CC Best Local Similarity 66.7%; Pred. No. 11;
 CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 7 RGRSGRGRGRG 18
 CC Db 325 RRGGRGRGRG 336
 CC
 CC RESULT 6
 CC FUS_BOVIN
 CC ID FUS_BOVIN STANDARD; PRT; 512 AA.
 CC AC Q28609;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE RNA-binding protein FUS (Pispen protein).
 CC GN FUS.
 CC OS Bos taurus (Bovine).
 CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC CC Bovidae; Bovinae; Bos.
 CC CC NCB1_TaxID=9913;
 CC RN NCB1 [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Heart;
 CC RX MEDLINE=96175600; PubMed=8631501;
 CC RA Allegro M.C., Allegro M.A.;
 CC RT "A nuclear protein regulated during the transition from active to
 CC RT quiescent phenotype in cultured endothelial cells";
 CC RT Dev. Biol. 174:288-297(1996).
 CC RN [2]
 CC RP CARBOHYDRATE-BINDING DOMAIN
 CC RX MEDLINE=20160719; PubMed=10694442;
 CC RA Allegro M.C.;
 CC RT "A C-terminal carbohydrate-binding domain in the endothelial cell
 CC RT regulatory protein, pigen: new function for an EWS family member";
 CC RL Exp. Cell Res. 255:270-277(2000).
 CC CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
 CC CC promotes ATP-independent annealing of complementary single-
 CC CC stranded DNAs and D-loop formation in superhelical double-stranded
 CC CC DNA. May play a role in maintenance of genomic integrity (By
 CC CC similarity).
 CC CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Binds SF1 (By
 CC CC similarity).
 CC CC -1- SUBCELLULAR LOCATION: Nuclear; exhibits diffuse staining
 CC CC throughout (excluding nucleoli), together with a small number of
 CC CC intensely stained focal points, or granules, and punctate staining
 CC CC along the nuclear envelope.
 CC CC -1- DOMAIN: The C-terminal domain binds carbohydrates.

```

CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U26024; AAC13543.1; -.
DR HSSP; P09651; 1HA1.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR001876; Znf_RanGDP.
DR Pfam; PF00076; rrm; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00547; ZNF_RBZ; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS01358; zf_RANBP2_1; 1.
DR PROSITE; PS50199; zf_RANBP2_2; 1.
DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KM Metal-binding.
FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.
FT DOMAIN 165 253 GLY-RICH.
FT DOMAIN 271 357 RNA-BINDING (RRM).
FT DOMAIN 357 512 ARG/GLY-RICH.
FT ZN_FING 408 439 RANBP2-TYPE.
SQ SEQUENCE 512 AA; 52240 MW; 3652329C044F1386 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 512;
Best Local Similarity 66.7%; Pred. NO. 18;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSXGRGGRG 18
Db 228 RGRGGRGGRG 239

RESULT 7
FUS_HUMAN
ID FUS_MOUSE STANDARD; PRT; 518 AA.
AC P56959;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE RNA-binding protein FUS (Fusgen protein).
GN FUS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=22829580; PubMed=12950080;
RA Alappat S.R., Zhang M., Zhao X., Allegro M.A., Allegro M.C.,
RA Burdall C.A.;
RT "Mouse p150 encodes a nuclear protein whose expression is
RT developmentally regulated during craniofacial morphogenesis.";
RL Dev. Dyn. 228:59-71(2003).
RN [2]
CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
CC promotes ATP-independent annealing of complementary single-
CC stranded DNAs and D-loop formation in superhelical double-stranded
CC DNA. May play a role in maintenance of genomic integrity (By
CC similarity).
CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts
CC with ILF3 and SPT (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

```

```

CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF224264; AAF70602.1; -.
DR HSSP; P09651; 1HA1.
DR MCD; MGT135363; Fus.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR001876; Znf_RanGDP.
DR Pfam; PF00076; rrm; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00547; ZNF_RBZ; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS01358; zf_RANBP2_1; 1.
DR PROSITE; PS50199; zf_RANBP2_2; 1.
DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KM Metal-binding.
FT DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.
FT DOMAIN 168 265 GLY-RICH.
FT DOMAIN 278 364 RNA-BINDING (RRM).
FT DOMAIN 364 518 ARG/GLY-RICH.
FT ZN_FING 415 446 RANBP2-TYPE.
SQ SEQUENCE 518 AA; 52673 MW; E06F231BFED78D6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 518;
Best Local Similarity 66.7%; Pred. NO. 18;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 RGRSXGRGGRG 18
Db 235 RGRGGRGGRG 246

RESULT 8
FUS_HUMAN
ID FUS_HUMAN STANDARD; PRT; 526 AA.
AC P35637;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE RNA-binding protein FUS (Oncogene FUS) (Translocated in
DE liposarcoma protein) (POMP75) (75 kDa DNA-pairing protein).
GN FUS OR TLS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93288139; PubMed=8510758;
RA Crozat A., Aman P., Mandahl N., Ron D.;
RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid
RT liposarcoma.";
RL Nature 363:640-644(1993).
RN [2]
CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
CC promotes ATP-independent annealing of complementary single-
CC stranded DNAs and D-loop formation in superhelical double-stranded
CC DNA. May play a role in maintenance of genomic integrity (By
CC similarity).
CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts
CC with ILF3 and SPT (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

```

RA Munakata N., Ohki M.;
 RT "Genomic structure of the human RBP56/hRAF168 and FUS/TLS genes.";
 RL Gene 221:191-198(1998).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Lung; and Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Ranev U., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, AND
 RX CHARACTERIZATION.
 RA MEDLINE=20036580; PubMed=10567410;
 RT Bechtold H., Kuroda M., Sok J., Ron D., Lopez B.S., Akhmedov A.T.;
 RT "Human 75-Kda DNA-pairing protein is identical to the pro-oncoprotein
 TLS/FUS and is able to promote D-loop formation.";
 RL J. Biol. Chem. 274:34337-34342(1999).
 RN [6]
 RP SEQUENCE OF 265-276; 317-331 AND 335-357, AND IDENTIFICATION.
 RX MEDLINE=99369251; PubMed=10442642;
 RT Bertrand P., Akhmedov A.T., Delacoste F., Durrbach A., Lopez B.S.;
 RT "Human POMP/5 is identified as the pro-oncogene TLF/FUS: both
 POMP/5 and POMP100 DNA homologous pairing activities are associated
 RT to cell proliferation.";
 RL Oncogene 18:4515-4521(1999).
 RN [7]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE=94243799; PubMed=8187069;
 RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;
 RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human
 RT myeloid leukemia with t(16;21) chromosomal translocation.";
 RL Cancer Res. 54:2865-2868(1994).
 RN [8]
 RP INTERACTION WITH SP1.
 RX MEDLINE=98325009; PubMed=9660765;
 RA Zhang D., Paley A.J., Childs G.;
 RT "The transcriptional repressor ZFM1 interacts with and modulates the
 RT ability of SWS to activate transcription.";
 RL J. Biol. Chem. 273:18086-18091(1998).
 CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and
 CC promotes ATP-independent annealing of complementary single-
 CC stranded DNAs and D-loop formation in superhelical double-stranded
 CC DNA. May play a role in maintenance of genomic integrity.
 CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts
 CC with ILF3 and SP1.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Bvent=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P35637-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P35637-2; Sequence=VSP_005798;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DISEASE: Involved in a form of malignant myeloid liposarcoma
 CC through a chromosomal translocation t(12;16)(q13;p11) that

CC INVOLVES DDIT3 and FUS.
 CC -1- DISEASE: Involved in a form of acute myeloid leukemia (AML)
 CC through a chromosomal translocation t(16;21)(p11;q22) that
 CC involves FUS and ERG.
 CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ch/announce/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL: S62140; AAB27102.1; -;
 CC EMBL: S62138; AAB27103.1; ALT_SEQ.
 CC EMBL: X71427; CA50558.1; ALT_SEQ.
 CC EMBL: X71428; CA50559.1; ALT_SEQ.
 CC EMBL: AF071213; AAC35285.1; -;
 CC EMBL: AF071213; AAC35284.1; -;
 CC EMBL: BC000402; AAH00402.1; -;
 CC EMBL: BC002459; AAH02459.1; -;
 CC PIR: S33799; S33799.
 CC HSSP: P09651; 1HA1.
 CC GeneW: HGNC:4010; FUS.
 CC GK: P35637; -;
 CC MIM: 137070; -;
 CC MIM: 151900; -;
 CC GO: GO:0005634; C:nucleus; TAS.
 CC GO: GO:0003723; F:RNA binding; TAS.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC InterPro: IPR001876; Znf_RanBP.
 CC Pfam: PF00076; rtm; 1.
 CC Pfam: PF00641; ZF_RanBP; 1.
 CC SMART: SMO0360; RRM; 1.
 CC SMART: SMO0547; ZNF_RBP2; 1.
 CC PROSITE: PS50102; RRM; 1.
 CC PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC PROSITE: PS01358; ZF_RANBP2_1; 1.
 CC PROSITE: PS01358; ZF_RANBP2_2; 1.
 CC KX Proco-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;
 CC KW Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;
 CC Metal-binding.
 CC FT DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.
 CC FT DOMAIN 166 267 GLY-RICH.
 CC FT DOMAIN 285 371 RNA-BINDING (RRM).
 CC FT DOMAIN 371 526 ARG/GLY-RICH.
 CC FT ZN_FING 422 453 RANBP2-TYPE.
 CC FT SITE 266 267 BREAKPOINT FOR TRANSLOCATION TO FORM
 CC FUS/TLS-CHOP ONCOGENE.
 CC FT VARSPLIC 64 65 TG -> S (in isoform short).
 CC FT CONFLICT 338 338 /FTid=VSP_005798.
 CC SQ SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;
 CC
 CC Query Match 46.7%; Score 43; DB 1; Length 526;
 CC Best Local Similarity 66.7%; Pred. No. 18;
 CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 7 RGRSYGRSGRSG 18
 CC Db 242 RGRGGRGGRG 253
 CC
 CC RESULT 9
 CC ID Y005_CABE1 STANDARD; PRT; 1232 AA.
 CC AC P34643;
 CC DT 01-FEB-1994 (Rel. 28, Created)
 CC DT 01-FEB-1994 (Rel. 28, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)

DR HSSP, Q27450; 1A33.
 DR MGD; MG1:97346; Nktr.
 DR InterPro; IPR002130; CSA_PPIase.
 DR Pfam; PF00160; pro_isomerase.1.
 DR PRINTS; PR00153; CSAPISMRASE.
 DR PROSITE; PS00170; CSA_PPIASE_1; 1.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 KM Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
 FT DOMAIN 1 176 PPIASE; CYCLOPHILIN-TYPE.
 FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
 FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
 FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
 FT DOMAIN 198 273 ARG/SER-RICH.
 FT DOMAIN 468 565 ARG/SER-RICH.
 FT DOMAIN 658 812 ARG/SER-RICH.
 FT DOMAIN 1303 1453 ARG-SER TANDEM REPEAT-RICH.
 SQ SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;

Query Match 46.7%; Score 43; DB 1; Length 1453;
 Best Local Similarity 81.8%; Pred. No. 54;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 RSXGRSGSGS 19
 Db 682 RSSGSGSGSGS 692

RESULT 12
 YDPC_SCHPO STANDARD; PRT; 150 AA.
 AC O14015;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C29A4.12c in chromosome I.
 GN SPAC29A4.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 CX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9712;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakert G., Aert R., Robben J., Gymnopoulos B.,
 RA Welfens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moesli D., Hilbert H.,
 RA Borzom K., Langer I., Beck A., Lehrach H., Reinhardt D., Pohl T.M.,
 RA Ezer P., Zimmermann W., Medler H., Wambutt R., Pirnalle B.,
 RA Goffeau A., Cadieu E., Drenth S., Gloux S., Jellare V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Spinkovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "the genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----

DR EMBL; Z97210; CAB10139.1; -.
 DR PIR; T38477; T38477.
 DR GeneDB; SPCmbe; SPAC29A4.12c; -.
 KW Hypothetical protein.
 SQ SEQUENCE 150 AA; 16781 MW; 3CFDCDE223106361 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 150;
 Best Local Similarity 56.2%; Pred. No. 6.8;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTRGRSXGRSGR 16
 Db 76 SPSSESTRNRSSGHTGR 91

RESULT 13
 FRZ2_DROME STANDARD; PRT; 694 AA.
 AC Q9VXX3; Q94916; Q9VXX2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fritzzled protein 2 precursor (Fritzzled-2) (dfz2).
 GN FZ2 OR CG9739/CG14083.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhant P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.,
 RT "A new member of the Fritzzled family from Drosophila functions as a
 RT Wingless receptor";
 RL Nature 382:225-230(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu D., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Docson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.T., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenisnson J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Iasko F., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod K.P., McPherson D.,
 RA Mervulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins, inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeletons of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage
 CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC -1- DOMAIN: Lys-Thr-X-X-TRP motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The FZ domain is involved in binding with Wnt ligands.
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U65589; AAC47273.1; -.
 DR EMBL: AE003518; AAF49185.2; -.
 DR PIR: S71786; S71786.
 DR FLYBase: FBgn0016797; fz2.
 DR GO: GO:0016021; C: integral to membrane, NAS.
 DR GO: GO:0017147; F: Wnt-protein binding, IDA.
 DR GO: GO:0006928; P: cell motility, IMP.
 DR GO: GO:0007163; P: establishment and/or maintenance of cell po. .; ISS.
 DR GO: GO:0008585; P: female gonad development, IMP.
 DR GO: GO:0016055; P: Wnt receptor signaling pathway, IDA.
 DR InterPro: IPR000539; Frizzled.
 DR InterPro: IPR000024; Fz domain.
 DR InterPro: IPR000832; GPCR secretin.
 DR Pfam: PF01534; Frizzled_1.
 DR Pfam: PF01392; Fz_1.
 DR PRINTS: PR00489; FRIZZLED.
 DR SMART: SM00063; FRI.1.
 DR PROSITE: PS50038; FZ.1.
 DR PROSITE: PS50261; G_PROTEIN_RECEP_F2_4.1.
 DR Multigene family: Receptor; G-protein coupled receptor; Transmembrane;
 DR Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.

FT	SIGNAL	1	22	POTENTIAL.
FT	CHAIN	23	694	FRIZZLED PROTEIN 2.
FT	DOMAIN	23	315	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	316	336	1 (POTENTIAL).
FT	TRANSMEM	337	352	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	353	373	2 (POTENTIAL).
FT	DOMAIN	374	397	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	398	418	3 (POTENTIAL).
FT	DOMAIN	419	439	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	440	460	4 (POTENTIAL).
FT	DOMAIN	461	482	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	483	503	5 (POTENTIAL).
FT	DOMAIN	504	534	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	535	555	6 (POTENTIAL).
FT	DOMAIN	556	584	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	585	605	7 (POTENTIAL).
FT	DOMAIN	606	694	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	187	225	GLY-RICH.
FT	DOMAIN	59	180	FZ.
FT	SITE	608	613	LYS-THR-X-X-TRP MOTIF.
FT	SITE	682	694	PDZ-BINDING.
FT	CARBOHYD	78	78	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	288	288	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	55	55	V -> A (IN REF. 1).
FT	CONFLICT	417	417	S -> T (IN REF. 1).
SO	SEQUENCE	694 AA;	75451 MW;	6C510P13CBAPB096 CRC64;

Query Match
 Best Local Similarity 52.6%;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

DB 1 SHOESTRGRSGSGSGSGS 19
 182 SYTEAGSGSGSGSGSGS 200

RESULT 14
 FRO_CRESP ID FRO_CRESP STANDARD; PRT; 1015 AA.
 AC 000586;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Frequency clock protein.
 GN FRO.
 OS *Creopus spinulosus* (Chromocorea spinulosa).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Hypocreales; Hypocrea.
 OX NCBI_Taxid=110619;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FGSC 4896;
 RX MEDLINE=97188515; PubMed=9037100;
 RA Lewis M.T., Morgan L.W., Feldman J.F.;
 RT "Analysis of frequency (frq) clock gene homologs: evidence for a
 RT helix-turn-helix transcription factor.";
 RL Mol. Gen. Genet. 253:401-414(1997).
 CC -1- FUNCTION: Circadian clock component involved in the generation of
 CC biological rhythms, in particular in rhythm stability, period
 CC length, and temperature compensation. Behaves as a negative
 CC element in circadian transcriptional loop (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FRO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U25850; AAA88072.1; -.
 DR PIR: T42013; T42013.


```

KW Biological rhythms; Transcription regulation; Nuclear protein.
FT DOMAIN 240 245 POLY-SER.
FT DOMAIN 356 368 POLY-GLN.
FT DOMAIN 443 451 POLY-PRO.
FT DOMAIN 584 588 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 892 913 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 1015 AA; 110972 MW; EA49E732ED741481 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 53;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
Db NHRKQKTGHSTGDSGSGN 601

RESULT 15
SKI2_YEAST STANDARD; PRT; 1287 AA.
AC P35207; Q06047;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Antiviral protein SKI2.
GN SKI2 OR YLR198C OR L8084.17.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93309467; PubMed=8321235;
RA Widner W.R., Wickner R.B.;
RT "Evidence that the Ski antiviral system of Saccharomyces cerevisiae
acts by blocking expression of viral mRNA.";
RL Mol. Cell. Biol. 13:4331-4341(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Uomanson M., Hillier L., Riles L., Albertmann K., Andre B., Ansoorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Eutlian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.V., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Ruer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe W., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrastarazu L.A., Vandenbol M., Verhaeselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambut R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
RL "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RT Nature 387:87-90(1997).
RN [3]
RP SEQUENCE OF 1-162 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95116323; PubMed=7816623;
RA Lygerou Z., Conesa C., Lesage P., Swanson R.N., Ruet A., Carlson M.,
RA Sentenac A., Seraphin B.;
RT "The yeast BDF1 gene encodes a transcription factor involved in the
expression of a broad class of genes including snRNAs.";
RL Nucleic Acids Res. 22:5332-5340(1994).
CC -1- FUNCTION: REPRESSIONS DSRNA VIRUS PROPAGATION BY SPECIFICALLY
BLOCKING TRANSLATION OF VIRAL MRNAs, PERHAPS RECOGNIZING THE
ABSENCE OF CAP OR POLY(A). ESSENTIAL FOR CELL GROWTH ONLY IN THE
PRESENCE OF M1 REPLICON. SEEMS TO BE A HELICASE.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: Belongs to the helicase family. SKI2 subfamily.
CC -----
CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13469; AAA35049.1; -
DR EMBL: U19729; AAB82356.1; -
DR EMBL: Z18944; CAA79378.1; -
DR PIR: S55954; S55954.
DR Geronline; 142461; -
DR SGD: S0004390; SKI2.
DR GO: GO:0005737; Cytoplasm; IDA.
DR GO: GO:0006402; P:RNA catabolism; IMP.
DR GO: GO:0006445; P:regulation of translation; IGI.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
KW Antiviral; Transcription regulation; RNA-binding; DNA-binding; Helicase;
KW ATP-binding.
FT NP BIND 351 358 ATP (POTENTIAL).
FT SITE 444 447 DEVH BOX.
FT DOMAIN 556 577 RNA-BINDING RG3-BOX (BY SIMILARITY).
FT DOMAIN 555 597 ARG/GLY-RICH.
FT CONFLICT 326 326 W -> C (IN REF. 1).
FT CONFLICT 759 760 QM -> L (IN REF. 1).
SQ SEQUENCE 1287 AA; 146058 MW; 7CCD36CFCDF8C32 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 1287;
Best Local Similarity 47.4%; Pred. No. 68;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
Db 550 SKTDNGRGSTRARGRGS 568

```

Search completed: September 28, 2004, 06:05:47
Job time : 8.22396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 / Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHQESTGRSGXGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_rhbc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	98.9	798	4 Q9H4U3	Q9H4U3 homo sapien
2	91	98.9	1084	4 Q01212	Q01212 homo sapien
3	90	97.8	797	4 Q16824	Q16824 homo sapien
4	90	97.8	990	4 Q15206	Q15206 homo sapien
5	90	97.8	1218	4 Q05331	Q05331 homo sapien
6	82	89.1	591	4 Q01720	Q01720 homo sapien
7	82	89.1	687	4 Q9H4U2	Q9H4U2 homo sapien
8	77	83.7	465	4 Q03838	Q03838 homo sapien
9	76	82.6	322	4 Q75370	Q75370 homo sapien
10	54	58.7	471	5 Q9UAY0	Q9UAY0 caenorhabdi
11	50	54.3	800	5 Q15845	Q15845 leishmania
12	50	54.3	1284	5 Q9V9Y3	Q9V9Y3 drosophila
13	48	52.2	822	3 Q9P312	Q9P312 neurospora
14	48	52.2	1229	3 Q86ZHI	Q86ZHI neurospora
15	47	51.1	52	16 Q98BE2	Q98BE2 rhizobium 1
16	47	51.1	349	16 Q7UD67	Q7UD67 rhodospirillum

17	47	51.1	436	16 Q82C67	Q82C67 streptomyces
18	46.5	50.5	813	12 Q91IN8	Q91IN8 hepatitis b
19	46	50.0	197	5 Q17626	Q17626 caenorhabdi
20	46	50.0	203	5 P91026	P91026 caenorhabdi
21	46	50.0	306	5 Q45386	Q45386 caenorhabdi
22	46	50.0	373	10 Q23876	Q23876 oryza sativ
23	46	50.0	399	4 Q9NTA9	Q9NTA9 homo sapien
24	46	50.0	467	4 Q9BWI8	Q9BWI8 homo sapien
25	46	50.0	737	4 Q9BO39	Q9BO39 homo sapien
26	46	50.0	836	10 Q9SKZ9	Q9SKZ9 arabidopsis
27	46	50.0	1007	10 Q8VYR8	Q8VYR8 arabidopsis
28	46	50.0	1105	10 Q9M383	Q9M383 arabidopsis
29	45.5	49.5	573	5 Q81GR2	Q81GR2 drosophila
30	45.5	49.5	1910	5 Q9V768	Q9V768 drosophila
31	45	48.9	243	10 Q9AX45	Q9AX45 oryza sativ
32	45	48.9	374	5 Q76561	Q76561 herpessvirus
33	45	48.9	476	12 Q80890	Q80890 herpessvirus
34	44.5	48.4	319	10 Q94A70	Q94A70 arabidopsis
35	44	47.8	89	16 Q9A9V2	Q9A9V2 caulobacter
36	44	47.8	229	11 Q9JW93	Q9JW93 mus musculus
37	44	47.8	259	4 Q86YA2	Q86YA2 mus musculus
38	44	47.8	282	3 Q871R2	Q871R2 neurospora
39	44	47.8	333	10 Q9SE88	Q9SE88 brassica o1
40	44	47.8	344	4 Q8N444	Q8N444 homo sapien
41	44	47.8	426	4 Q8V76	Q8V76 homo sapien
42	44	47.8	588	5 Q819N5	Q819N5 schistosoma
43	44	47.8	669	11 Q8BRU5	Q8BRU5 mus musculus
44	44	47.8	734	11 Q99M09	Q99M09 mus musculus
45	44	47.8	867	5 Q18218	Q18218 caenorhabdi

ALIGNMENTS

RESULT 1
Q9H4U3 PRELIMINARY; PRT; 798 AA.
AC Q9H4U3
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJL4N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.
RL Submitted (OCF-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; ..
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGRIN.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
Query Match 98.9%; Score 91; DB 4; Length 798;
Best Local Similarity 94.7%; Pred. No. 9.5e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHQESTGRSGXGRSGSGS 19
DB 428 SHQESTGRSGXGRSGSGS 446
RESULT 2
Q01212 PRELIMINARY; PRT; 1084 AA.
AC Q01212; Q01840; (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Created)

```

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60501; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:Intermediate filament; NAS.
DR GO; GO:0005198; F:Structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
FT NON TER
FT SEQUENCE 1084 AA; 115271 MW; 80C4640B9D5A362D CRC64;
SQ
Query Match 98.9%; Score 91; DB 4; Length 1084;
Best Local Similarity 94.7%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTRGRSGRSGSGS 19
Db 64 SH0ESTRGRSGRSGSGS 82

RESULT 3
ID Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN Flg.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60501; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:Intermediate filament; NAS.
DR GO; GO:0005198; F:Structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
FT NON TER
FT SEQUENCE 797 AA; 85176 MW; 60B6184763BDA86B CRC64;
SQ
Query Match 97.8%; Score 90; DB 4; Length 797;
Best Local Similarity 94.7%; Pred. No. 1.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTRGRSGRSGSGS 19
Db 427 SH0ESTRGRSGRSGSGS 445

```

```

RESULT 4
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN Flg.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60494; AAA63244.1; -.
DR EMBL; M60494; AAA63244.1; JOINED.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGRIN.
FT NON TER
FT SEQUENCE 990 AA; 106453 MW; A8396F10FA91991 CRC64;
SQ
Query Match 97.8%; Score 90; DB 4; Length 990;
Best Local Similarity 94.7%; Pred. No. 1.8e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTRGRSGRSGSGS 19
Db 875 SH0ESTRGRSGRSGSGS 893

RESULT 5
ID Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGRIN (PROFILAGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.C., Markov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC - FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFIDE-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC - P.TM: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF

```

CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; M96943; AAA36487.1; -.
 DR PIR; A48118; A48118.
 DR HSSP; P02593; 1CDM.
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0005509; F:calcium ion binding; TAS.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 6.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
 KM Polymorphism.
 FT CA_BIND 19 32 SITE I (BY SIMILARITY).
 FT CA_BIND 62 73 SITE II (BY SIMILARITY).
 FT NON_TER 1218 1218
 SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 1218;
 Best Local Similarity 94.7%; Pred. No. 2.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGSRXGSRGSGS 19
 Db 1097 SHQESTRGSRGSRGSGS 1115

RESULT 6
 Q01720 ID 001720 PRELIMINARY; PRT; 591 AA.
 AC 001720;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Presland R.B., Haydock P.V., Fleckman P., Niyonskisi W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 RT organization and identification of an S-100-like calcium binding
 RT domain at the amino terminus";
 RL J. Biol. Chem. 267:23772-23781(1992).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; L01089; AAA60177.1; -.
 DR EMBL; L01090; AAA60176.1; -.

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KM Polymorphism.
 FT PROPEP 1 293 POTENTIAL.
 FT CHAIN 294 467 FILAGGRIN.
 FT PROPEP 468 474 POTENTIAL.
 FT CHAIN 475 >591 FILAGGRIN.
 FT CA_BIND 19 32 SITE I (BY SIMILARITY).
 FT CA_BIND 62 73 SITE II (BY SIMILARITY).
 FT NON_TER 591 591
 SQ SEQUENCE 591 AA; 66366 MW; 381491625C75B369 CRC64;

Query Match 89.1%; Score 82; DB 4; Length 591;
 Best Local Similarity 89.5%; Pred. No. 2.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTRGSRXGSRGSGS 19
 Db 449 SHQESTRGSRGSRGSGS 467

RESULT 7
 Q094U2 ID 094U2 PRELIMINARY; PRT; 687 AA.
 AC 094U2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laid G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 CC EMBL; AL356504; CAC13172.1; -.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CAP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON_TER 687 687
 SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 89.1%; Score 82; DB 4; Length 687;
 Best Local Similarity 89.5%; Pred. No. 2.7e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19
 DB 449 SHOESTRGRSGRSGSGS 467

RESULT 8

OQ3838 PRELIMINARY; PRT; 465 AA.

ID OQ3838;

AC 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

FLG. FILAGRIN (PROFILAGRIN) (Fragment).

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RE MEDLINE=91064347; PubMed=2248957;

RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin

gene.",

RL Biochemistry 29:9432-9440(1990).

RN (2)

RP REVISIONS.

RE MEDLINE=91255199; PubMed=2043621;

RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin

gene.",

RL Biochemistry 30:5814-5814(1991).

-1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND

PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

-1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND

WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

REPEATS

-1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,

INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN

GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

PROTEOLYTICALLY CLEAVED.

EMBL: M60499; AAA63246.1; -

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam: PF03516; Filaggrin; 3.

DR PRINTS: PR00487; FILAGRIN.

FT NON_TER 1 1

FT 465 465

SEQUENCE 465 AA; 50280 MW; C683744C5E134097 CRC64;

Query Match 83.7%; Score 77; DB 4; Length 465;

Best Local Similarity 84.2%; Pred. No. 0.00012;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 19

DB 227 SHOESTRGRSGRSGSGS 245

RESULT 9

OQ3838 PRELIMINARY; PRT; 322 AA.

ID OQ3838;

AC 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

FLG. Epi-dermal filaggrin (Fragment).

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN (1)

RP SEQUENCE FROM N.A.

RE MEDLINE=99101527; PubMed=9886436;

RA Girbal-Munauter E., Durieux J.-J., Arnaud M., Dalpon P., Sebba M.,

RA Vincent C., Simon M., Sensu T., Masson-Bessiere C.,

RA Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated

antifilaggrin autoantibodies are posttranslationally generated on

RT various sites of (pro)filaggrin by deamination of arginine residues.";

RL J. Immunol. 162:585-594(1999).

DR EMBL: AF043380; AAC23559.1; -

DR GO: GO:0005198; P:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam: PF03516; Filaggrin; 2.

DR PRINTS: PR00487; FILAGRIN.

FT NON_TER 1 1

FT 322 322

SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match 82.6%; Score 76; DB 4; Length 322;

Best Local Similarity 83.3%; Pred. No. 0.00012;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSGRSGSGS 18

DB 305 SHOESTRGRSGRSGSGS 322

RESULT 10

OQ3838 PRELIMINARY; PRT; 471 AA.

ID OQ3838;

AC 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

FLG. W03G1.5 protein.

GN W03G1.5.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Felodermidae; Caenorhabditis.

OX NCBI_Taxid=6239;

RN (1)

RP SEQUENCE FROM N.A.

RE MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berke M.,

RA Bonfield U., Burton J., Connell M., Copey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,

RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thiermy-Mieg U., Thomas K., Vaudin X., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat U., Wohldman P.,

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans.";

RL Nature 368:32-38(1994).

RN (2)

RP SEQUENCE FROM N.A.

RE STRAIN=BRISTOL N2;

RA Pauley A., Scheet P., Harper M.;

RT "The sequence of C. elegans cosmid W03G1.";

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

RN (3)

RP SEQUENCE FROM N.A.

RE STRAIN=BRISTOL N2;

RA Waterston R.;

RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF125964; AAD14753.1; -

DR PIR: T33997; T33997.

DR WormBep: W03G1.5; CBI7283.

```

SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A985B CRC64;
Query Match 58.7%; Score 54; DB 5; Length 471;
Best Local Similarity 57.9%; Pred. No. 0.92;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSGS 19
   : ||| ||| |||
Db 167 SSRSPRGRGRGRGGRGSGS 185

RESULT 11
ID 015845 PRELIMINARY; PRT; 800 AA.
AC 015845;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE L1439.4.
GN L1439.4.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
CX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=9178987; Pubmed=10077609;
RA Wyler P.J., Audleman L., Devos T., Hixson G., Kiser P., Lemley C.,
RA Magnus C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AB001274; AAC24675.1; -.
DR PIR; T02852; T02852.
SQ SEQUENCE 800 AA; 83782 MW; DA57AE627933CBAC CRC64;

QY 2 HOESTRGSRGSRGSGS 19
   : ||| ||| |||
Db 429 HRDVGRLSTTASGRSGS 446

Query Match 54.3%; Score 50; DB 5; Length 800;
Best Local Similarity 55.6%; Pred. No. 7.8;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

RESULT 12
ID Q9V9Y3 PRELIMINARY; PRT; 1284 AA.
AC Q9V9Y3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CG11339-PA.
GN CG11339.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richardson S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,
RA Brendon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gaber G.L.,
RA Abilil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Balise R.M., Baer P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

```

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.W., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Miska S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celinker S.E.,
RA Ciamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kromoller B., Marshall B., Milburn G.F., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnick F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Flybase;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Flybase;
RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB003777; AAF57145.2; -.
DR Flybase; FBgn003841; CG11339.
DR GO; GO:0005856; C:cytoskeleton; IEA.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_41; 1.
DR PRINTS; PR00395; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00507; FERM_3; 1.
SQ SEQUENCE 1284 AA; 136966 MW; EDF4119838A9A6447 CRC64;

Query Match 54.3%; Score 50; DB 5; Length 1284;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGSRGSG 18
   : ||| ||| |||
Db 545 SNESELSRGSRGSRGSG 562

RESULT 13
ID Q9P312 PRELIMINARY; PRT; 822 AA.
AC Q9P312;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

```

```

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12P1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algen V., Hohlseil J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -.
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR005120; Smg_4_UFP3.
DR Pfam; PF000076; rtm; 1.
DR Pfam; PF03467; Smg4_UFP3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 52.2%; Score 48; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTRGSGXGRSGSGS 19
DB 414 RESAGRRGRGRGCGT 430

RESULT 14
Q86ZHI PRELIMINARY; PRT; 1229 AA.
AC Q86ZHI;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Related to cell division cycle 2-related protein kinase 7.
GN 7P4.060.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algen V., Hohlseil J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX294020; CAD70910.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0000910; P:cytokinesis; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; kinase; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00228; S_TKC; 1.
DR SMART; SM00219; TYKC; 1.

```

```

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Cell division; Kinase.
SQ SEQUENCE 1229 AA; 136776 MW; 43BE61656FF73E4 CRC64;

Query Match 52.2%; Score 48; DB 3; Length 1229;
Best Local Similarity 57.9%; Pred. No. 27;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTRGSGXGRSGSGS 19
DB 242 SHHERRSRDRGRSGRSRS 260

RESULT 15
Q98BE2 PRELIMINARY; PRT; 52 AA.
AC Q98BE2;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical protein ms15615.
GN MS15615.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP003007; BAB52030.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 52 AA; 5836 MW; 505124C0543255PA CRC64;

Query Match 51.1%; Score 47; DB 16; Length 52;
Best Local Similarity 64.3%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRGSGXGRSGSGS 19
DB 39 TRAGSGRAGTSGS 52

```

Search completed: September 28, 2004, 06:12:41
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHQESTRGSRXGSRGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_290Jan04:*
1: geneseqp1960s:*
2: geneseqp1950s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	97.8	19	2	AAW61507	AAW61507 Peptide c
2	90	97.8	19	2	AAW61514	AAW61514 Peptide c
3	90	97.8	1467	5	ABB97605	ABB97605 Novel hum
4	86	93.5	19	2	AAW61517	AAW61517 Peptide c
5	85	92.4	19	2	AAW61516	AAW61516 Peptide c
6	84	91.3	19	2	AAW61506	AAW61506 Peptide c
7	84	91.3	19	2	AAW61508	AAW61508 Peptide c
8	84	91.3	19	2	AAW61515	AAW61515 Peptide c
9	84	91.3	19	2	AAW61509	AAW61509 Peptide c
10	84	91.3	19	2	AAW61511	AAW61511 Peptide c
11	84	91.3	19	2	AAW61505	AAW61505 Peptide c
12	78	84.8	19	2	AAW61512	AAW61512 Peptide c
13	78	84.8	19	2	AAW61513	AAW61513 Peptide c
14	78	84.8	19	2	AAW61510	AAW61510 Peptide c
15	75	81.5	21	2	AAW61520	AAW61520 Peptide X
16	75	81.5	22	4	AAE07235	AAE07235 ICP1546 p
17	66	71.7	330	2	AAV22956	AAV22956 Human fil
18	66	71.7	330	2	AAV22955	AAV22955 Human fil
19	66	71.7	330	2	AAV22957	AAV22957 Human fil
20	62	67.4	330	2	AAV22954	AAV22954 Human fil
21	57	62.0	477	6	ABO07142	ABO07142 Novel hum
22	54	58.7	1711	4	AAW79819	AAW79819 Human pro
23	54	58.7	1951	4	AAW78835	AAW78835 Human pro
24	50	54.3	1199	4	ABBS5274	ABBS5274 Drosophila
25	47	51.1	184	3	AAW57041	AAW57041 Human pro

26	46	50.0	41	4	AAW16780	AAW16780 Peptide #
27	46	50.0	41	4	ABW35764	ABW35764 Peptide #
28	46	50.0	41	4	AAW29265	AAW29265 Peptide #
29	46	50.0	41	4	ABW30600	ABW30600 Peptide #
30	46	50.0	41	4	ABW21188	ABW21188 Protein #
31	46	50.0	41	4	AAW68954	AAW68954 Human bon
32	46	50.0	41	4	AAW56574	AAW56574 Human bra
33	46	50.0	41	4	ABG50617	ABG50617 Human liv
34	46	50.0	41	4	AAW04497	AAW04497 Peptide #
35	46	50.0	41	5	ABG38540	ABG38540 Human pep
36	46	50.0	443	4	AAW40069	AAW40069 Human pol
37	46	50.0	574	4	AAW41855	AAW41855 Human pol
38	46	50.0	752	4	AAU23535	AAU23535 Novel hum
39	45.5	49.5	1945	4	ABW64947	ABW64947 Drosophila
40	44	47.8	229	7	ABW59825	ABW59825 Rat Prote
41	44	47.8	229	7	ABW59833	ABW59833 Rat Prote
42	44	47.8	229	7	ABW59837	ABW59837 Rat Prote
43	44	47.8	229	7	ABW59829	ABW59829 Rat Prote
44	44	47.8	1038	7	ADCC03412	ADCC03412 Rice flow
45	44	47.8	1308	6	ABW57072	ABW57072 Human tau

ALIGNMENTS

RESULT 1	
AAW61507	
ID	AAW61507 standard; peptide; 19 AA.
XX	
AC	AAW61507;
XX	
DT	26-OCT-1998 (first entry)
DE	Peptide cfc3, based on cDNA of a profilaggrin repeat.
XX	
XX	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW	solid phase synthesis; peptide amide; polyclonal antibody;
KM	monoclonal antibody.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site 11
FT	/note="Citruiline"
XX	
PN	W09822503-A2.
XX	
PD	28-MAY-1998.
XX	
PF	14-NOV-1997; 97MO-NL000624.
XX	
PR	15-NOV-1996; 96NL-01004539.
XX	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.
XX	
PI	Van Venrooij WJW, Schellekens GA, Raats JWH, Hoet RMA;
XX	
DR	WPI; 1998-398613/34.
XX	
PT	Peptide derived from an antigen recognised by autoantibodies - is
PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT	used in diagnosis of the disease.
XX	
PS	Disclosure; Page 6; 199p; English.
XX	
CC	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC	the profilaggrin antigen which is recognised by autoantibodies from
CC	patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC	The peptides were created by using standard solid phase synthesis, which
CC	produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGS 19
 |||||
 DB 1 SHOESTRGRSKXGRSGS 19

RESULT 2
 AAW61514
 ID AAW61514 standard; peptide; 19 AA.

XX AAW61514;

DT 26-OCT-1998 (first entry)

DE Peptide c.f. based on cDNA of a profilaggrin repeat.

XX Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX WO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-396613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61500 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.6e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGS 19
 |||||
 DB 1 SHOESTRGRSKXGRSGS 19

RESULT 3

ABB97605
 ID ABB97605 standard; protein; 1467 AA.

XX ABB97605;

XX 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

XX Human; anti-neoplastic; anti-inflammatory; immunomodulator;
 KW anti-infective; cerebroprotective; cytoprotective; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

XX Homo sapiens.

XX WO200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

XX 11-SEP-2000; 2000US-00659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AV, Yang Y, Weinman T, Drmanac RT;

XX WPI; 2002-292408/33.

XX N-PSDB; ABN32791.

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate haemostasis or thrombolysis e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention

SQ Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.7e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSKXGRSGS 19
 |||||
 DB 1097 SHOESTRGRSKXGRSGS 1115

RESULT 4
 AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

DT 26-OCT-1998 (first entry)

DE Peptide c.f. based on cDNA of a profilaggrin repeat.

XX Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

```

OS Synthetic.
OS Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 93.5%; Score 86; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.1e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHOESTGRSGRSGRSGS 19
   |||||:|||||||
Db 1 SHOESTGRSGRSGRSGS 19

RESULT 5
AAM61516
ID AAM61516 standard; peptide; 19 AA.
XX
XX AAM61516;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfe, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;

```

```

XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 92.4%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTGRSGRSGRSGS 19
   |||||:|||||||
Db 1 SHOESTGRSGRSGRSGS 19

RESULT 6
AAM61506
ID AAM61506 standard; peptide; 19 AA.
XX
XX AAM61506;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 9
FT FT /note="Citruiline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 7
 ID AAM61508 standard; peptide; 19 AA.

XX AAM61508;

DT 26-OCT-1998 (first entry)

DE Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

FT Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "Citrulline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI, 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

1 Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 8
 ID AAM61515 standard; peptide; 19 AA.

XX AAM61515;

DT 26-OCT-1998 (first entry)

DE Peptide cfa, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.

OS Synthetic.
 OS Homo sapiens.

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI, 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRXGRSGSGS 19
 DB 1 SHOESTRGSRXGRSGSGS 19

RESULT 9
 ID AAM61509 standard; peptide; 19 AA.

XX AAM61509;

DT 26-OCT-1998 (first entry)

XX

DE Peptide cfc5, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 16
 FT /note= "Citruiline"
 PN WO9822503-A2.
 XX
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHQESTGRSGRSGRSGS 19
 DB 1 SHQESTGRSGRSGRSGS 19

RESULT 10
 AAM61511
 ID AAM61511 standard; peptide; 19 AA.
 XX
 AC AAM61511;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX
 XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 7

FT /note= "Citruiline"
 FT Modified-site 11
 FT /note= "Citruiline"
 PN WO9822503-A2.
 XX
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 PS Disclosure; Page 6; 19pp; English.
 XX
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHQESTGRSGRSGRSGS 19
 DB 1 SHQESTGRSGRSGRSGS 19

RESULT 11
 AAM61505
 ID AAM61505 standard; peptide; 19 AA.
 XX
 AC AAM61505;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX
 XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruiline"
 PN WO9822503-A2.
 XX
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 AC |||||
 AC 1 SHQESTXGRSGRSGSGS 19
 DB

RESULT 12
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 AC AAW61512;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 7
 FT Modified-site /note= "Citnulline"
 FT Modified-site 13
 FT /note= "Citnulline"
 FT
 XX WO9822503-A2.
 PN
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 AC |||||
 AC 1 SHQESTXGRSGRSGSGS 19
 DB

RESULT 13
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.
 XX
 AC AAW61513;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc9, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 7
 FT Modified-site /note= "Citnulline"
 FT Modified-site 16
 FT /note= "Citnulline"
 FT
 XX WO9822503-A2.
 PN
 XX 28-MAY-1998.
 PD
 XX 14-NOV-1997; 97WO-NL000624.
 PF
 XX 15-NOV-1996; 96NL-01004539.
 PR
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
 |||||
 1 SHOESTXGRSGRSGRSGS 19

RESULT 14
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

AC AAW61510;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfcd6, based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT Modified-site 9 /note= "Citruiline"
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.
 XX
 PD 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 2e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGRSGS 19
 |||||
 1 SHOESTXGRSGRSGRSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide XI based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruiline"
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.
 XX
 PD 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI
 DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 6.6e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HOESTGRSGRSGRSGS 19
 |||||
 4 HOESTXGRSGRSGRSGS 21

Search completed: September 28, 2004, 06:24:12
 Job time : 55.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-3
Perfect score: 92
Sequence: 1 SHQESTRGSRXGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/ECTS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-3
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-4
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-7
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-8
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	57	62.0	477	15	US-10-161-927-62	Sequence 62, Appl
17	50	54.3	133	16	US-10-437-963-136367	Sequence 136367,
18	49	53.3	123	16	US-10-767-701-32436	Sequence 32436, A
19	48	52.2	179	12	US-10-425-114-35345	Sequence 5345, A
20	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
21	47	51.1	436	14	US-10-156-763-13022	Sequence 13022, A
22	47	51.1	51.1	16	US-10-437-963-188040	Sequence 188040,
23	47	51.1	563	16	US-10-437-963-198539	Sequence 198539,
24	47	51.1	889	16	US-10-437-963-181020	Sequence 181020,
25	46	50.0	41	9	US-09-864-761-36486	Sequence 36486, A
26	46	50.0	164	16	US-10-437-963-112419	Sequence 112419,
27	46	50.0	21.1	14	US-10-002-631C-68	Sequence 268, App
28	46	50.0	373	16	US-10-437-963-143403	Sequence 143403,
29	46	50.0	46.1	16	US-10-437-963-185969	Sequence 185969,
30	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
31	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
32	45	48.9	227	16	US-10-437-963-181049	Sequence 181049,
33	45	48.9	243	16	US-10-437-963-103033	Sequence 103033,
34	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
35	44	47.8	78	16	US-10-437-963-203073	Sequence 203073,
36	44	47.8	155	16	US-10-437-963-161956	Sequence 161956,
37	44	47.8	445	12	US-10-424-599-267922	Sequence 267922,
38	44	47.8	612	12	US-10-424-599-230819	Sequence 230819,
39	44	47.8	1071	16	US-10-437-963-118389	Sequence 118389,
40	44	47.8	1270	16	US-10-408-765A-4253	Sequence 2253, Ap
41	44	47.8	1321	16	US-10-271-507-2	Sequence 2, Appl
42	43.5	47.3	19723	15	US-10-084-846A-5	Sequence 5, Appl
43	43	46.7	26	13	US-10-056-407-22	Sequence 22, Appl
44	43	46.7	46	13	US-10-056-407-20	Sequence 20, Appl
45	43	46.7	50	14	US-10-029-386-27692	Sequence 27692, A

ALIGNMENTS

RESULT 1
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michel Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2, 4e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 2
US-09-308-150-11
; Sequence 11, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-11

Query Match      97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 3
US-09-308-150-14
; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-14

Query Match      97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 4
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-13

Query Match      92.4%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 5
US-09-308-150-1
; Sequence 1, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-1

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-14

Query Match      93.5%; Score 86; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 9.1e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 4
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-13

Query Match      92.4%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHOESTRGRSGRSGSGS 19
      |||||
Db      1 SHOESTRGRSGRSGSGS 19

RESULT 5
US-09-308-150-1
; Sequence 1, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
US-09-308-150-1

```

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

RESULT 6
US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederlandsche Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

RESULT 7
US-09-308-150-4

Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederlandsche Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSRGSGS 19
Db 1 SHOESTRGSRGSRGSGS 19

RESULT 8
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederlandsche Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 9

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.7%; Pred. No. 1.8e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 10

US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 11

US-09-308-150-6

Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 1 SHOESTRGRSGRSGSGS 19

RESULT 12

US-09-308-150-8

Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHQSTGRSGXGRSGSGS 19
Db 1 SHQSTGRSGXGRSGSGS 19

```

```

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 SHQSTGRSGXGRSGSGS 19
Db 1 SHQSTGRSGXGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00039;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 HQSTGRSGXGRSGSGS 19
Db 4 HQSTGRSGXGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.003INPUS00 INNS:031
; CURRENT APPLICATION NUMBER: US/09/747,029A
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 22
; LENGTH: 22
; TYPE: PRT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-3

Perfect score: 92

Sequence: 1 SHQSTRGRSGRSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	49	53.3	629	US-09-252-991A-22901	Sequence 22901, A
2	47	51.1	406	US-09-252-991A-19857	Sequence 19857, A
3	45	48.9	395	US-09-252-991A-21702	Sequence 21702, A
4	45	48.9	778	US-09-252-991A-18605	Sequence 18605, A
5	44	47.8	369	US-09-252-991A-20245	Sequence 20245, A
6	44	47.8	374	US-09-252-991A-33040	Sequence 33040, A
7	44	47.8	483	US-09-252-991A-19015	Sequence 19015, A
8	44	47.8	518	US-09-252-991A-25967	Sequence 25967, A
9	43	46.7	26	US-09-252-991A-25967	Sequence 25967, A
10	43	46.7	46	US-09-252-991A-27107	Sequence 27107, A
11	43	46.7	159	US-09-252-991A-27107	Sequence 27107, A
12	43	46.7	653	US-09-513-057C-13	Sequence 13, Appl
13	42	45.7	85	US-09-252-991A-29339	Sequence 29339, A
14	42	45.7	105	US-09-252-991A-32394	Sequence 32394, A
15	42	45.7	175	US-09-252-991A-21718	Sequence 21718, A
16	42	45.7	803	US-09-252-991A-23614	Sequence 23614, A
17	42	45.7	2237	US-08-354-973-1	Sequence 1, Appl
18	41.5	45.1	237	US-09-252-991A-17927	Sequence 17927, A
19	41.5	45.1	398	US-09-252-991A-19488	Sequence 19488, A
20	41.5	45.1	415	US-09-328-352-6430	Sequence 6430, Ap
21	41	44.6	120	US-09-252-991A-17025	Sequence 17025, A
22	41	44.6	133	US-09-252-991A-17025	Sequence 17025, A
23	41	44.6	177	US-09-252-991A-30896	Sequence 30896, A
24	41	44.6	204	US-09-252-991A-21317	Sequence 21317, A
25	41	44.6	226	US-09-252-991A-22052	Sequence 22052, A
26	41	44.6	255	US-09-252-991A-19598	Sequence 19598, A
27	41	44.6	404	US-09-489-039A-11606	Sequence 11606, A

28	41	44.6	478	US-09-252-991A-25411	Sequence 25411, A
29	41	44.6	482	US-09-252-991A-16654	Sequence 16654, A
30	41	44.6	686	US-09-252-991A-20509	Sequence 20509, A
31	40.5	44.0	96	US-09-252-991A-19216	Sequence 19216, A
32	40	43.5	134	US-09-252-991A-23430	Sequence 23430, A
33	40	43.5	210	US-09-134-001C-4065	Sequence 4065, Ap
34	40	43.5	213	US-09-252-991A-32093	Sequence 32093, A
35	40	43.5	237	US-09-252-991A-33040	Sequence 33040, A
36	40	43.5	256	US-09-252-991A-26244	Sequence 26244, A
37	40	43.5	421	US-09-252-991A-32336	Sequence 32336, A
38	40	43.5	442	US-09-252-991A-23554	Sequence 23554, A
39	40	43.5	646	US-09-328-352-6017	Sequence 6017, Ap
40	40	43.5	722	US-09-984-890-4	Sequence 4, Appl
41	40	43.5	724	US-09-984-890-2	Sequence 2, Appl
42	40	43.5	745	US-09-523-849-36	Sequence 36, Appl
43	40	43.5	877	US-09-252-991A-25547	Sequence 25547, A
44	40	43.5	908	US-08-356-354-6	Sequence 6, Appl
45	40	43.5	908	US-08-778-656-6	Sequence 6, Appl

ALIGNMENTS

```

RESULT 1
US-09-252-991A-22901
; Sequence 22901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22901
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22901

Query Match      53.3%; Score 49; DB 4; Length 629;
Best Local Similarity 52.9%; Pred. No. 5.6;
Matches 9; Mismatches 4; Indels 0; Gaps 0;

QY      3 QESTGRSGRSGRSGS 19
DB      272 RQSHGRAGRAGRHGS 288

RESULT 2
US-09-252-991A-19857
; Sequence 19857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19857
; LENGTH: 406
; TYPE: PRT

```

ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-19857

Query Match	51.1%;	Score 47;	DB 4;	Length 406;
Best Local Similarity	50.0%;	Pred. No. 7.4;		
Matches	9;	Conservative	2;	Mismatches 7;
				Indels

```
QY      2 HQESTRGRSXGRSGRGS 19
         | : | | | | | | :
Db     127 HGRSRGRDRAGRGRHGA 144
```

RESULT 3
US-09-252-991A-21702

```

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; DATES: 09/06/88

```

Query Match	48.9%;	Score 45;	DB 4;	Length 395;
-------------	--------	-----------	-------	-------------

```
QY      1 SHQESTRGRSXGRSGRG 18
          ||| : |||
Db      251 SHQGASEGRGDHESGRPG 268
```

RESULT 4
US-09-252-991A-18605
; Sequence 18605, Application US/09252991A

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

Query Match	48.9%	Score	45	DB	4	Length	778
Best Local Similarity	58.8%	Pred. No.	30				
Matches	10	Conservative	2	Mismatches	5	Indels	0
						Gaps	0

```
QY      3 QESTRGRSXGRSGSGS 19
          | : ||||| | ||| :
Db      11 QTATRGRSEQRRERSGN 27
```

RESULT 5
US-09-252-991A-20245
; Sequence 20245, Application US/09252991A

? APPLICANT: Marc J. Rubenfield et al.
 ? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ? TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ? FILE REFERENCE: 107396.136
 ? CURRENT APPLICATION NUMBER: US/09/252,991A

Query Match	47.8%;	Score 44;	DB 4;	Length 369;
Best Local Similarity	47.1%;	Pred. NO. 20;		
Matches	8;	Conservative	2;	Mismatches 7; Indels 0; Gaps 0;

```
QY      2 HQESTRGRSXGRSGRG 18
      | : ||| : | |||
Db      221 HRRTRGSAAGGQGRHG 23
```

RESULT 6
US-09-252-991A-33040
; Sequence 33040, Application US/09252991A

APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

Query Match	47.8%;	Score 44;	DB 4;	Length 374;
Best Local Similarity	61.5%;	Pred. No. 20;		
Matches	8;	Conservative	2;	Mismatches 3; Indels 0; Gaps 0;

QY	4	ESTRGRSXGRSGR	16
		: :	
Db	258	ETLRGRTAGRQGR	270

RESULT 7
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A

: APPLICANT: Marc J. Rubenstein et al.
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : TITLE OF INVENTION: PSEUDOMONAS FOR DIAGNOSTICS AND THERAPEUTICS
 : FILE REFERENCE: 107196.136
 : CURRENT APPLICATION NUMBER: US/09/252,991A


```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-19015

```

```

Query Match      47.8%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY      2 H0E8TRGRSXGRSG 18
Db      94 HAEQDPGAGTGRAGRAG 110

```

```

RESULT 8
US-09-252-991A-25967
; Sequence 25967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25967
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25967

```

```

Query Match      47.8%; Score 44; DB 4; Length 518;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      8 GRGXGRSGRG 18
Db      415 GRGSRGSRG 425

```

```

RESULT 9
US-09-297-981-22
; Sequence 22, Application US/09297981
; Patent No. 6362007
; GENERAL INFORMATION:
; APPLICANT: MEHENS, LYDIE
; APPLICANT: LUHMANN, REINHARD GEORGE
; APPLICANT: UNION, ANN
; APPLICANT: RAYNACKERS, JOSEPH
; TITLE OF INVENTION: METHYLATED, SMD HOMOLOGOUS PEPTIDES, REACTIVE WITH THE
; TITLE OF INVENTION: ANTIBODIES FROM SERA OF LIVING BEINGS AFFECTED WITH
; TITLE OF INVENTION: SYSTEMIC LUPUS ERYTHEMATOSUS
; FILE REFERENCE: INNS011--
; CURRENT APPLICATION NUMBER: US/09/297,981
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 22
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-297-981-22

```

```

Query Match      46.7%; Score 43; DB 4; Length 26;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY      7 RGRSGRGSG 18
Db      7 RGRGGGRGSG 18

```

```

RESULT 10
US-09-297-981-20
; Sequence 20, Application US/09297981
; Patent No. 6362007
; GENERAL INFORMATION:
; APPLICANT: MEHENS, LYDIE
; APPLICANT: LUHMANN, REINHARD GEORGE
; APPLICANT: UNION, ANN
; APPLICANT: RAYNACKERS, JOSEPH
; TITLE OF INVENTION: METHYLATED, SMD HOMOLOGOUS PEPTIDES, REACTIVE WITH THE
; TITLE OF INVENTION: ANTIBODIES FROM SERA OF LIVING BEINGS AFFECTED WITH
; FILE REFERENCE: INNS011--
; CURRENT APPLICATION NUMBER: US/09/297,981
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 20
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-297-981-20

```

```

Query Match      46.7%; Score 43; DB 4; Length 46;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY      7 RGRSGRGSG 18
Db      27 RGRGGGRGSG 38

```

```

RESULT 11
US-09-252-991A-27107
; Sequence 27107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27107
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27107

```

```

Query Match      46.7%; Score 43; DB 4; Length 159;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 SHOESTRGRSXGRSGRS 19
| | | | | | | | | |
DB 5 SASSTRGRSAMRSTASS 23

RESULT 12
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 46.7%; Score 43; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 51;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTRGRSXGRSGRS 19
| | | | | | | | | |
DB 573 SRQVSTASASGREGISGS 591

RESULT 13
US-09-252-991A-29339
; Sequence 29339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29339
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29339

Query Match 45.7%; Score 42; DB 4; Length 85;
Best Local Similarity 80.0%; Pred. No. 9;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GRSGXGRSGRS 17
| | | | | | | | | |
DB 73 GRAAGRSGRS 82

RESULT 14
US-09-252-991A-32394
; Sequence 32394, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32394
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32394

Query Match 45.7%; Score 42; DB 4; Length 105;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 TRGRSXGRSGRSG 18
| | | | | | | | | |
DB 65 TRRCAGRCGATG 77

RESULT 15
US-09-252-991A-21788
; Sequence 21788, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21788
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21788

Query Match 45.7%; Score 42; DB 4; Length 175;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHOESTRGRSXGRSGR 16
| | | | | | | | | |
DB 62 SRRPRAGRAAGRTGR 77

Search completed: September 28, 2004, 06:26:42
Job time: 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-4
Perfect score: 92
Sequence: 1 SHOESTRGRSGSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	2248	2 A35938	profilaggrin - hum
2	84	91.3	416	2 A32947	profilaggrin precursor
3	75	81.5	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2M47	E2 protein - human
5	50	54.3	800	2 T02852	probable membrane
6	48	52.2	772	2 T27512	hypothetical prote
7	48	52.2	822	2 T51049	related to nucleol
8	47	51.1	471	2 T33997	hypothetical prote
9	46	50.0	373	2 T02976	probable DNA bindi
10	46	50.0	836	2 G84727	probable DNA topoi
11	45	50.0	1232	2 S40766	hypothetical prote
12	46	50.0	1829	2 T35681	probable sensory h
13	45	49.5	135	2 A46338	RNA-binding protei
14	44.5	48.4	553	2 T27245	hypothetical prote
15	44.5	48.4	568	2 H88904	protein Y57G11C.9
16	44.5	48.4	659	2 T27246	hypothetical prote
17	44	47.8	229	2 T07219	nuclear protein SR
18	44	47.8	776	2 A87330	nuclear protein histi
19	44	47.8	790	2 T05576	hypothetical prote
20	44	47.8	867	2 T27136	hypothetical prote
21	44	47.8	871	2 T27135	hypothetical prote
22	44	47.8	2420	2 A84652	hypothetical prote
23	43.5	47.3	463	2 T51194	hypothetical prote
24	43.5	47.3	2561	2 T24864	hypothetical prote
25	43	46.7	123	2 T16234	hypothetical prote
26	43	46.7	150	2 B71185	probable ribosoma
27	43	46.7	151	2 F75145	lsu ribosomal prot
28	43	46.7	210	2 C84404	hypothetical prote
29	43	46.7	521	2 B86332	hypothetical prote

30	43	46.7	638	2 I53169	cytochrome 2 - hu
31	43	46.7	1507	2 B47328	natural killer cel
32	42	45.7	231	2 U70463	nuclear factor I -
33	42	45.7	306	2 T21220	hypothetical prote
34	42	45.7	350	2 A40459	nuclear phosphopro
35	42	45.7	369	2 F96788	protein T4012.22 l
36	42	45.7	374	2 A37282	52K active chroma
37	42	45.7	538	2 H86335	T20H2.2 protein -
38	42	45.7	694	2 S71786	wingless receptor
39	42	45.7	849	2 A96592	hypothetical prote
40	42	45.7	1015	2 T42013	frequency clock pr
41	42	45.7	1337	2 T30291	dextranase - Strep
42	41	45.1	483	2 S36470	E2 protein - human
43	41	44.6	136	2 T35632	probable transpos
44	41	44.6	174	2 A27293	probable Y chromos
45	41	44.6	176	2 F71370	probable single-st

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C>Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gam, S.Q.; Mcaride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A>Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; PMID:91064347; PMID:2246957
A/Accession: A35938
A>Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-References: GB:J02929
C/Genetics:
A/Gene: GDB:FLG
A/Cross-References: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:46-569/Region: flaggrin repeat
F:570-893/Region: flaggrin repeat
F:1074-1397/Region: flaggrin repeat
F:1573-1896/Region: flaggrin repeat

Query Match 97.8%; Score 90; DB 2; Length 2248;
Best Local Similarity 94.7%; Pred. No. 8.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGSGSGS 19
DB 227 SHOESTRGRSGSGSGS 245

RESULT 2
A32947
profilaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/Mckinley-Grant, D.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A>Title: Characterization of a cDNA clone encoding human flaggrin and localization of cl
A/Reference number: A32947; PMID:89296901; PMID:2740331
A/Accession: A32947
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-References: GB:M2435; NID:g182604; PIDN:AAA52454.1; PID:g182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

probable sensory histidine kinase - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T35681
C/Harris, D.; Taylor, K.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1998
A:Reference number: Z21587

A:Accession: T35681

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1829 <HAR>

A:Cross-references: EMBL:AL031031; PTDN:CAA19849.1; GSPDB:GN00070; SCOEDB:SC7C7.03

A:Experimental source: strain A3(2)

A:Genetics:

A:Gene: SCOEDB:SC7C7.03

Query Match 50.0%; Score 46; DB 2; Length 1829;
Best Local Similarity 60.0%; Pred. No. 49;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 STRGRSGXSGRSGS 19
|:|||||
21 SNQGRSRGCGARSGS 35

RESULT 13

A:Accession: A46398

A:Status: preliminary

A:Molecule type: RNA-binding protein 1 - fruit fly (Drosophila melanogaster)

A:Residues: 1-135 <KIM>

A:Cross-references: GB:L04929; NID:G158223; PTDN:AAA28850.1; PID:G158224

A:Note: authors translated the codon GCT for residue 65 as Arg

A:Gene: FlyBase:Rbp1

A:Cross-references: FlyBase:FBgn0010252

A:Introns: 106/2

A:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein F12-74/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 49.5%; Score 45.5; DB 2; Length 135;
Best Local Similarity 52.2%; Pred. No. 4.7;

Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 4 ESTGRSR-----GXSGRSGS 19
|:|||||

DB 80 EMSGSRDRRRGEGSSGSGS 102

RESULT 14

T27245

A:Accession: T27245

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-553 <WIL>

A:Cross-references: EMBL:Z99281; PTDN:CAB54457.1; GSPDB:GN00022; CESP:Y57G11C.9a

A:Experimental source: clone Y57G11C

A:Genetics:

A:Gene: CESP:Y57G11C.9a

A:Map position: 4

A:Introns: 67/2; 109/2; 382/1; 418/2; 447/1; 496/3

Query Match 48.4%; Score 44.5; DB 2; Length 553;

Best Local Similarity 64.7%; Pred. No. 26;
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 CESTGRSRGXSGRSGS 19
|:|||||

DB 218 RSNTRSRK-SGGRSGS 233

RESULT 15

H88904

A:Accession: H88904

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <STO>

A:Cross-references: GB:chr_IV; PTDN:CAB16531.1; PID:G3881206; GSPDB:GN00022; CESP:Y57G11C.9

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Map position: 4

Query Match 48.4%; Score 44.5; DB 2; Length 568;
Best Local Similarity 64.7%; Pred. No. 27;

Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 CESTGRSRGXSGRSGS 19
|:|||||

DB 233 RSNTRSRK-SGGRSGS 248

Search completed: September 28, 2004, 06:14:59

Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136,952 Million cell updates/sec

Title: US-09-308-150-4

Sequence: 1 SHQSTRGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	91.3	416	1	FLIIA_HUMAN
2	58	63.0	506	1	VE2_HPVA7
3	48	52.2	772	1	WR11_CAEEL
4	46	50.0	963	1	Y144_HUMAN
5	46	50.0	1232	1	Y005_CAEEL
6	45.5	49.5	135	1	RBP1_DROME
7	44	47.8	124	1	SMD1_DROME
8	43	46.7	120	1	LSM4_CAEEL
9	43	46.7	150	1	RL19_PYRHO
10	43	46.7	151	1	RL19_PYRAB
11	43	46.7	210	1	YO43_HUMAN
12	43	46.7	638	1	K220_HUMAN
13	43	46.7	1453	1	NKCR_MOUSE
14	43	45.7	231	1	NP1L_PIG
15	42	45.7	375	1	SR55_DROME
16	42	45.7	694	1	FRZ2_DROME
17	42	45.7	1015	1	FRQ_CRESP
18	42	45.7	1180	1	TYK2_MOUSE
19	42	45.7	1337	1	DEXT_STRDO
20	41.5	45.1	388	1	ROG_MOUSE
21	41.5	45.1	391	1	ROG_HUMAN
22	41.5	45.1	483	1	VE2_HPVA4
23	41	44.6	174	1	Y119_HUMAN
24	41	44.6	176	1	SSB_TREPA
25	41	44.6	205	1	GARI_YEAST
26	41	44.6	414	1	NSR1_YEAST
27	41	44.6	454	1	VE2_HPVA5
28	41	44.6	502	1	VE2_HPVA7
29	41	44.6	503	1	VE2_HPVA6
30	41	44.6	509	1	VE2_HPVA6
31	41	44.6	713	1	DDX4_RAT
32	41	44.6	825	1	SE5_RAT
33	41	44.6	2404	1	SON_MOUSE

ALIGNMENTS

34	41	44.6	2426	1	SON_HUMAN	P18583 homo sapien
35	40.5	44.0	1509	1	MYSN_ACACA	P05659 acanthamoeb
36	40	43.5	150	1	YDPC_SCHPO	O14015 schizosacch
37	40	43.5	356	1	DEP3_MOUSE	P58269 mus musculu
38	40	43.5	526	1	2ABA_YEAST	O00462 saccharomyc
39	40	43.5	546	1	GHT5_SCHPO	P78831 schizosacch
40	40	43.5	659	1	HMT1_DROME	P22807 drosophila
41	40	43.5	677	1	SG1_MOUSE	P16014 mus musculu
42	40	43.5	730	1	DPOL_HPVA4	P12933 hepatitis b
43	40	43.5	842	1	DPOL_HPVA4	P31870 hepatitis b
44	40	43.5	843	1	DPOL_HPVA4	P03157 hepatitis b
45	40	43.5	1053	1	SPS_SOLTU	Q43845 solanum tub

RESULT 1

AC	FLIIA_HUMAN	STANDARD;	PRT;	416 AA.
AC	P20930;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Flaggrin precursor (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89296901; PubMed=2740331;			
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,			
RA	Carnizero L., Croce C.W., Huebner K., Lessin S.R., Stehert P.M.,			
RT	"Characterization of a cDNA clone encoding human flaggrin and			
RT	localization of the gene to chromosome region 1q21."			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).			
RN	[2]			
RP	CITRULLINATION.			
RX	MEDLINE=96374388; PubMed=8780679;			
RA	Senhhu T., Kan S., Ogawa H., Manabe W., Asaga H.,			
RT	"Preferential deimination of keratin Ki and flaggrin during the			
RT	terminal differentiation of human epidermis."			
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).			
CC	-I- FUNCTION: Aggregates keratin intermediate filaments and promotes			
CC	disulfide-bond formation among the intermediate filaments during			
CC	terminal differentiation of mammalian epidermis.			
CC	-I- PFM: Flaggrin is initially synthesized as a large, insoluble,			
CC	highly phosphorylated precursor containing many tandem copies of			
CC	324 AA, which are not separated by "large linker". The precursor			
CC	is deposited as keratohyalin granules. During terminal			
CC	differentiation it is dephosphorylated and proteolytically			
CC	cleaved.			
CC	-I- PFM: Undergoes deimination of some arginine residues			
CC	(citrullination).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL collaboration			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by, and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M24355; AA52454.1; -			
CC	PIR; A32947; A32947.			
CC	Genew; HGNC:3748; FLG.			
CC	-----			
CC	MM; M135940; -			
CC	GO; GO:0005882; C:intermediate filament; NMS.			
CC	GO; GO:0005198; F:structural molecule activity; NMS.			
CC	GO; GO:0007275; P:development; NMS.			
CC	InterPro; IPR003303; Flaggrin.			

DR Pfam, PF03516; Flaggrin; 2.
 DR PRINTS; PR00487; FLAGGRIN.
 KW Phosphorylation; Citrullination; Developmental protein.
 FT NON TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
 Best Local Similarity 89.5%; Pred. No. 4,4e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGSGS 19
 Db 7 SHOESTRGRSGXSGSGS 25

RESULT 2

VE2-HPV47 STANDARD; PRT; 506 AA.
 ID VE2-HPV47
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2.
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 RT type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405 (1990).
 CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION
 CC IT BINDS TO THE E2B RESPONSE ELEMENT (5'-ACNNNNNGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2B'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STRICTLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M32305; AAA46979.1; -
 CC PIR; D53324; W2WT47.
 DR HSSP; P03122; 2BOP.
 DR InterPro; IPR000427; E2_C.
 DR InterPro; IPR001866; E2_N.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF00511; E2_C; 1.
 DR Pfam; PF00506; E2_N; 1.
 DR Pfam; PF000672; E2_C; 1.
 DR Pfam; PF000678; E2_N; 1.
 DR ProDom; PD000672; E2_C; 1.
 DR ProDom; PD000678; E2_N; 1.
 KW Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BFB725065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.07;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGSGS 19

Db 342 SREGNTRGRGRGRGSGS 360

RESULT 3

MR1 CAEEL STANDARD; PRT; 772 AA.
 ID MR1 CAEEL
 AC Q23255;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Double-strand break repair protein mre-11.
 GN MRE-11 OR ZC302.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
 RP GLU-139.
 RX MEDLINE=21135651; PubMed=11238374;
 RA Chin G.M., Villeneuve A.M.;
 RT "C. elegans mre-11 is required for meiotic recombination and DNA
 RT repair but is dispensable for the meiotic G12 DNA damage
 RT checkpoint.";
 RL Genes Dev. 15:522-534 (2001).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2.
 RA Kelly P.F.;
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
 CC processing.
 CC -1- FUNCTION: Required for meiotic crossing over and chiasma
 CC formation. Pachytene morphology and homolog pairing are normal.
 CC Vital in long term for maintenance of reproductive capacity of
 CC subsequent generations.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the MRE11/RAD32 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z73978; CAA98292.1; -
 CC PIR; T27512; T27512.
 DR GenMOnline; 208612; -
 DR WormPep; ZC302.1; CE06573.
 DR InterPro; IPR003701; DNA_repair.
 DR InterPro; IPR004643; M-Pdestrase.
 DR InterPro; IPR007281; Mre11_DNA_bind.
 DR Pfam; PF00149; MetAllopos; 1.
 DR Pfam; PF04152; Mre11_DNA_bind; 1.
 DR TIGRFam; TIGR00583; mre11; 1.
 KW DNA repair; Hydrolase; Nuclease; Endonuclease; Exonuclease;
 KW Nuclear protein; Manganese; Meiosis.
 FT MUTAGEN 139 139 E->K: IN MRE-11-ME41; DEFECTIVE IN
 FT MEIOTIC CHROMOSOME DEGRADATION.
 SQ SEQUENCE 772 AA; 86913 MW; 74A0754C2604C4B CRC64;

Query Match 52.2%; Score 48; DB 1; Length 772;
 Best Local Similarity 52.6%; Pred. No. 4.2;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 1 SHOESTRGRSGXSGSGS 19


```

Db      697 SKOPTTRGRGRGARGAGAS 715

RESULT 4
Y144 HUMAN
ID      Y144 HUMAN      STANDARD;      PRT;      983 AA.
AC      014157;
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Hypothetical protein KIAA0144.
GN      KIAA0144.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Bone marrow;
RX      MEDLINE=96127530; PubMed=8590280;
RA      Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT      "Prediction of the coding sequences of unidentified human genes. IV.
RT      The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT      analysis of cDNA clones from human cell line KG-1.";
RL      DNA Res. 2:167-174(1995).
CC      -1- SIMILARITY: Contains 1 UBA domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL, D63478; BAA09765.1; -.
DR      InterPro: IPR000449; UBA_domain.
DR      Pfam: PF00627; UBA; 1.
DR      SMART, SM00165; UBA; 1.
DR      PROSITE, PSS0030; UBA; 1.
KM      Hypothetical protein. 89
FT      DOMAIN 49
SQ      SEQUENCE 983 AA; 103930 MW; 50579CFA293EEA68 CRC64;

Query Match      50.0%; Score 46; DB 1; Length 983;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy      7 RGRSRGSGRSG 18
Db      177 RGRGRGSGRRG 188

RESULT 5
Y005 CAEEL
ID      Y005 CAEEL      STANDARD;      PRT;      1232 AA.
AC      P34643;
DT      01-FEB-1994 (Rel. 28, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical protein ZK512.5 in chromosome III.
GN      ZK512.5.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RX      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA      Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

```

```

RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA      Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA      Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA      Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callahan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA      Sims M., Smaildon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA      Sulten J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA      Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA      Woldman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans.";
RL      Nature 368:32-38(1994).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL, 222177; CAA80146.1; -.
DR      PIR, S40766; S40766.
DR      WormPeP, ZK512.5; CE00411.
KM      Hypothetical protein.
SQ      SEQUENCE 1232 AA; 134923 MW; 6DFC35D664A8D6A CRC64;

Query Match      50.0%; Score 46; DB 1; Length 1232;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      2 HOESTRGRSGXSG 15
Db      110 HONSSRGPSGSPSG 123

RESULT 6
RBP1 DROME
ID      RBP1 DROME      STANDARD;      PRT;      135 AA.
AC      Q02427; Q26271; Q9VGM8;
DT      15-DEC-1998 (Rel. 37, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      RNA-binding protein 1.
GN      RBP1 OR CG37136.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX      MEDLINE=94040720; PubMed=1340470;
RA      Kim Y.-C., Zuo P., Manley J.L., Baker B.S.;
RT      "The Drosophila RNA-binding protein RBP1 is localized to
RT      transcriptionally active sites of chromosomes and shows a functional
RT      similarity to human splicing factor ASF/SF2.";
RL      Genes Dev. 6:2569-2579(1992).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RC      STRAIN=Berkley;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blaise R.G., Change M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Bens P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA      Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

```

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mentrulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Palazonzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 14-52 FROM N.A.
 RX MEDLINE=93109300; PubMed=8417324;
 RA Kim Y.J., Baker B.S.;
 RT "Isolation of Ruv-type RNA-binding protein genes and the analysis of
 RL their relatedness by using a numerical approach,"
 RN Mol. Cell. Biol. 13:174-183(1993).
 RN [4]
 RP FUNCTION:
 RX MEDLINE=95393975; PubMed=7664738;
 RA Heinrichs V., Baker B.S.;
 RT "The *Drosophila* SR protein RBP1 contributes to the regulation of
 RT doublesex alternative splicing by recognizing RBP1 RNA target
 RT sequences,"
 RL EMBL J. 14:3987-4000(1995).
 CC -1- FUNCTION: Contributes to the activation of female-specific DSX
 CC splicing in vivo by recognizing the RBP1 target sequences within
 CC the purine-rich polypyrimidine tract of the female-specific 3'
 CC splice site.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RBP1-A;
 CC IsoId=Q02427-1; Sequence=Displayed;
 CC Name=2; Synonyms=RBP1-B;
 CC IsoId=Q02427-2; Sequence=VSP_005817;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
 CC -1- PTM: Extensively phosphorylated on serine residues in the RS
 CC domain (Probable).
 CC -1- SIMILARITY: Belongs to the splicing factor SR family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L04829, AAA28850.1, -
 DR EMBL, AE003688, AAF54555.1, -
 DR EMBL, S51691, AAB24622.1, -
 DR PIR, A46398, A46398.
 DR PIR, A48110, A48110.

DR FlyBase; FBgn010252; RBP1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PSS0030; RRM_RNP_1; FALSE_NEG.
 DR Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
 KW Phosphorylation.
 FT DOMAIN 11 84 RNA-BINDING (RRM).
 FT DOMAIN 82 135 ARG/SER-RICH (RS DOMAIN).
 FT VARSPLIC 107 135 Missing (in isoform 2).
 FT VARSPLIC /FTid=VSP_005817.
 FT CONFLICT 14 14 Y -> F (IN REF. 3).
 FT CONFLICT 65 65 R -> A (IN REF. 1).
 SQ SEQUENCE 135 AA; 15446 MW; 895DCE902518D991 CRC64;
 Query Match 49.5%; Score 45.5; DE 1; Length 135;
 Best local Similarity 52.2%; Pred. No. 1.6;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
 QY 4 ESTGSR-----GXGSRGSG 19
 DB 80 EMSGSRSDRRRGEGSGSRGSG 102
 RESULT 7
 SMD1_DROME STANDARD; PRT; 124 AA.
 ID SMD1_DROME Q9V002; Q9SYR6;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable small nuclear ribonucleoprotein Sm D1 (snRNP core protein D1)
 DE (Sm-D1)
 GN SNRPD9 OR ECDNA;R39488 OR C010753.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Calniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck B., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mentrulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,
 RA Palazonzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

```

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Moodie T., Morley S.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=2242606; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Garth H., Krommiller B., Pacleb J.W., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- FUNCTION: Essential for pre-mRNA splicing. Implicated in the
CC formation of stable, biologically active snRNP structures (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the snRNP core protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003540; AAF49893.1; -.
DR EMBL: AY071359; AAL48981.1; -.
DR FLYbase; FBgn0016940; snRNP69D.
DR GO; GO:0030532; C:small nuclear ribonucleoprotein complex; ISS.
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; ISS.
DR GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; ISS.
DR InterPro: IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR KMW: Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing.
FT DOMAIN 86 124 ARG/LYS-RICH (BASIC).
FT DOMAIN 100 118 9 X 2 AA TANDEM REPEATS OF R-G.
FT CONFLICT 52 52 P -> H (IN REF. 2).
SQ SEQUENCE 124 AA; 13796 MW; 36FA51D3FB6855 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 124;
Best Local Similarity 50.0%; Pred. No. 2.6;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 1 SHOSTGRSGRGGRG 18
Db 94 SGRVGNRGGRGGRG 111

```

```

CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20864; AAC46661.1; -.
DR PIR; T16234; T16234.
DR Wormpep; F32A5.7; CE01277.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR Prodom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
DR KMW: Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 3.7;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 3 QESTGRSGRGGRG 18
Db 93 REQSRGGRGGRGGRG 108

RESULT 9
RL19 PYRHO
AC 059437; STANDARD; PRT; 150 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L19B.
GN RPL19B OR PH1759.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCB1_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyana A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -1- SIMILARITY: Belongs to the L19B family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000007; BAA30873.1; -.

```

DR PIR; B71185; Ribosomal_L19e.
 DR InterPro: IPR00196; Ribosomal_L19e.
 DR Pfam; PF01280; Ribosomal_L19e; 1.
 DR PRODOM; PD004823; Ribosomal_L19e; 1.
 DR PROSITE; PS00526; RIBOSOMAL_L19E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 150 AA; 17997 MW; 412223A854825B69 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 150;
 Best Local Similarity 41.2%; Pred. No. 4.5;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSGXSGRSG 18
 DB 67 HEQKKGRHGRGSRKG 83

RESULT 10

RL19_PYPAB STANDARD; PRT; 151 AA.
 AC Q9V1V3;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 10-OCT-2003 (Rel. 42; Last annotation update)
 DE 50S ribosomal protein L19S.
 GN RPL19E OR PYPAB03230 OR PAB2134.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 CC NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RX MEDLINE=22511545; PubMed=12622808;
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
 RA Poch O., Priet D., Quelellou J., Ripp R., Thierry J.-C.,
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P;
 RT "An integrated analysis of the genome of the hyperthermophilic
 archaeon Pyrococcus abyssi."
 RL Mol. Microbiol. 47:1493-1512(2003).
 CC -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ248284; CAB49245.1; -
 DR PIR; F75145; F75145.
 DR InterPro; IPR00196; Ribosomal_L19e.
 DR Pfam; PF01280; Ribosomal_L19e; 1.
 DR PRODOM; PD004823; Ribosomal_L19e; 1.
 DR PROSITE; PS00526; RIBOSOMAL_L19E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 151 AA; 18187 MW; 85FC0805643437A0 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 151;
 Best Local Similarity 41.2%; Pred. No. 4.5;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTRGRSGXSGRSG 18
 DB 67 HEQKKGRHGRGSRKG 83

RESULT 11

YO43_HALN1 STANDARD; PRT; 210 AA.
 AC Q9HMH2;
 DT 10-OCT-2003 (Rel. 42; Created)

DT 10-OCT-2003 (Rel. 42; Last sequence update)
 DT 10-OCT-2003 (Rel. 42; Last annotation update)
 DE Hypothetical protein Vng2543c.
 GN VNG2543C.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 CC Halobacteriaceae; Halobacterium.
 CC NCBI_TaxID=64091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahatras G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
 RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Madocks D.G., Jablonski P.E., Krebs M.P., Angwine C.M., Dale H.,
 RA Isebacher T.A., Beck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- SIMILARITY: Contains 1 AMMECR1 domain.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE005130; AAG20599.1; -
 DR PIR; C84404; C84404.
 DR HAMAP; MF_00645; atypical; 1.
 DR InterPro; IPR002733; DUF51.
 DR Pfam; PF01871; AMMECR1; 1.
 DR PRODOM; PD009671; DUF51; 1.
 DR TIGRFA; TIGR00296; TIGR00296; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 210 AA; 22964 MW; D4CB30CEB307EFF CRC64;

Query Match 46.7%; Score 43; DB 1; Length 210;
 Best Local Similarity 61.1%; Pred. No. 6.5;
 Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 4 ESTRGRS-RGXSGRSGS 19
 DB 55 ESTRGRGRURCGAGHS 72

RESULT 12

K220_HUMAN STANDARD; PRT; 638 AA.
 ID K220_HUMAN
 AC Q01546;
 DT 01-JUN-1994 (Rel. 29; Created)
 DT 01-JUN-1994 (Rel. 29; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2P) (K2P)
 DE (CK 2P).
 GN KRT2P.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93114504; PubMed=1282112;
 RA Collin C., Onayoun J.P., Grund C., Franke W.W.;
 RA "Suprabasal marker proteins distinguishing keratinizing squamous
 RT epithelia: cyokeratin 2 polypeptides of oral masticatory epithelium
 RT and epidermis are different."
 RL Differentiation 51:137-148(1992).
 CC -1- FUNCTION: Probably contributes to terminal cornification.

```

CC CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC CC keratinocytes.
CC CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC CC microfilillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC CC (neutral to basic; 56-70 kDa) [K1 to K8].
CC CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M99063; AAA35746.1; -.
CC CC PIR; I53169; I53169.
CC CC GO; GO:0005882; C:intermediate filament; NAS.
CC CC GO; GO:0005198; F:structural molecule activity; NAS.
CC CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
CC CC InterPro; IPR001664; IF.
CC CC InterPro; IPR002957; Keratin_I.
CC CC InterPro; IPR003054; Keratin_II.
CC CC Pfam; PF00038; filament; 1.
CC CC PRINTS; PR01248; TYPE1KERATIN.
CC CC PRINTS; PR01276; TYPE2KERATIN.
CC CC PROSITE; PS00226; IF; 1.
CC CC Intermediate filament; Coiled coil; Keratin; Phosphorylation.
CC CC DOMAIN 1 182 HEAD.
CC CC FT DOMAIN 183 492 ROD.
CC CC FT DOMAIN 493 638 TAIL.
CC CC FT DOMAIN 183 218 COIL_1A.
CC CC FT DOMAIN 219 237 LINKER_1.
CC CC FT DOMAIN 238 329 COIL_1B.
CC CC FT DOMAIN 330 353 LINKER_12.
CC CC FT DOMAIN 354 492 COIL_2.
CC CC FT MOD_RS 57 57 PHOSPHORYLATION (BY SIMILARITY).
CC CC SQ SEQUENCE 638 AA; 65871 MW; 9B743AB8B72076AF CRC64;

Query Match 46.7%; Score 43; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTGRGRGSGRSGS 17
Db 7 KKSFGSGSGSGSGS 21

RESULT 13
NCRC_MOUSE STANDARD; PRT; 1453 AA.
ID NCRC_MOUSE
AC P30415;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NK-tumor recognition protein (Natural-Killer cells cyclophilin-
DE related protein) (NK-TR protein).
GN NKTR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93133824; PubMed=8421688;
RA Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
RA Oretallo J.R.;
RT "A cyclophilin-related protein involved in the function of natural
RT killer cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RP REVISIONS TO C-TERMINUS.

```

```

CC CC STRAIN=BA/B/c; TISSUE=Blood;
CC CC RA Anderson S.K.;
CC CC RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: Component of a putative tumor-recognition complex.
CC CC Involved in the function of NK cells.
CC CC -1- SIMILARITY: Contains 1 cyclophilin-like p1ase domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U04289; AAA37500.2; ALT_INIT.
CC CC DR HSSP; Q27450; IA33.
CC CC DR MGD; MGI:97346; Nktr.
CC CC DR InterPro; IPR002130; CSA_P1ase.
CC CC DR Pfam; PF00160; pro.isomerase; 1.
CC CC DR PRINTS; PR00170; CSA_P1ASE_1; 1.
CC CC DR PROSITE; PS00072; CSA_P1ASE_2; 1.
CC CC DR Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
CC CC KW DOMAIN 1 176 P1ASE, CYCLOPHILIN-TYPE.
CC CC FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
CC CC FT DOMAIN 198 273 ARG/SER-RICH.
CC CC FT DOMAIN 468 565 ARG/SER-RICH.
CC CC FT DOMAIN 658 812 ARG/SER-RICH.
CC CC FT DOMAIN 1303 1453 ARG-SER TANDDEM REPEAT-RICH.
CC CC SQ SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;

Query Match 46.7%; Score 43; DB 1; Length 1453;
Best Local Similarity 81.8%; Pred. No. 51;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 RSRGSGRSGS 19
Db 682 RSGSGSGRSGS 692

RESULT 14
ID NFIL_PIG STANDARD; PRT; 231 AA.
AC P14057;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nuclear factor 1 (NF-1) (CCAAT-box binding transcription factor) (CTF)
DE (TGCA-binding protein) (Fragment).
GN NF1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296883; PubMed=2841167;
RA Meisternest M., Rogge L., Donath C., Gander I., Lottspeich F.,
RA Meitz R., Dobner T., Roessler R., Stelzer G., Winnacker E.L.;
RT "Isolation and characterization of the porcine nuclear factor I (NF1)
RT gene.";
RL FEBS Lett. 236:27-32(1988).
RN [2]
RP PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89088257; PubMed=3207762;
RA Gander I., Roessler R., Rogge L., Meisternest M., Schneider R.,
RA Meitz R., Lottspeich F., Winnacker E.L.;
RT "Purification methods for the sequence-specific DNA-binding protein
RT nuclear factor I (NF1) -- generation of protein sequence

```

```

RT Information."
RL Biochim. Biophys. Acta 951:411-418(1988).
CC -1- FUNCTION: Recognizes and binds the palindromic sequence 5'-
CC TTGGCANNNGCCAA-3' present in viral and cellular promoters and in
CC the origin of replication of adenovirus type 2. These proteins are
CC individually capable of activating transcription and replication.
CC -1- SUBUNIT: Binds DNA as a homodimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=P14057-1; Sequence=Displayed;
CC -1- SIMILARITY: Belongs to the CTF/NF-1 family.
CC -1- SIMILARITY: Contains 1 DNA/MH1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X12764; CAA31254.1; -.
CC PIR, J70463; J70463.
CC TRANSFAC; T00173; -.
CC TRANSFAC; T00538; -.
CC InterPro; IPR000647; CTF_NF1.
CC InterPro; IPR003619; DwaRtn_A.
CC Pfam; PF03165; MH1; 1.
CC SMART; SMO0523; DWA; 1.
CC PROSITE; PS00349; CTF_NF1; 1.
CC Transcription regulation; DNA replication; DNA-binding; Activator;
CC Nuclear protein; Multigene family; Alternative splicing.
CC DOMAIN
CC 65 173 DWA.
CC NON_TER
CC 231 231
CC SEQUENCE 231 AA; 26162 MW; A463C8BEE3D0E743 CRC64;
SQ
Query Match 45.7%; Score 42; DB 1; Length 231;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 4 ESTRGRSRGXSGSG 18
DB 186 EAGRARGRGSDGREG 200
RESULT 15
SRSS_DROME STANDARD; PRT; 375 AA.
AC P26686; Q24252;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine-arginine protein 55 (SRP55) (Enhancer of deformed) (52-kDa
DE bracketing protein) (B52 protein).
GN E(DPD) OR SR55 OR R855 OR B52.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyrididae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14; 125-131 AND 136-147.
RC STRAIN=CL; TISSUE=Embryo;
RX MEDLINE=92011900; PubMed=1717489;
RA Roth M.B., Zahler A.M., Stoik J.A.;
RT "A conserved family of nuclear phosphoproteins localized to sites of
RT polymerase II transcription."
RL J. Cell Biol. 115:587-596(1991).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).

```

```

RC TISSUE=Embryo;
RX MEDLINE=91357476; PubMed=1885003;
RA Champin D.T., Frasch M., Saunweber H., Lis J.T.;
RT "Characterization of a Drosophila protein associated with boundaries
RT of transcriptionally active chromatin."
RL Gene Dev. 5:1611-1621(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=92159037; PubMed=1741384;
RA Mayeda A., Zahler A.M., Krainer A.R., Roth M.B.;
RT "Two members of a conserved family of nuclear phosphoproteins are
RT involved in pre-mRNA splicing."
RL Proc. Natl. Acad. Sci. U.S.A. 89:1301-1304(1992).
RN [4]
RP FUNCTION.
RX MEDLINE=95021280; PubMed=7935465;
RA Ring H.Z., Lis J.T.;
RT "The SR protein B52/SRP55 is essential for Drosophila development."
RL Mol. Cell. Biol. 14:7499-7506(1994).
CC -1- FUNCTION: Essential for development. May have a critical role in
CC splicing or in controlling alternative splice site use of at least
CC some pre-mRNA in vivo. Not required for all splicing. May play a
CC general role in the condensation or decondensation of chromatin.
CC -1- SUBCELLULAR LOCATION: Nuclear; associated with boundaries of
CC transcriptionally active chromatin.
CC -1- TRANSCRIPTIONALLY ACTIVE CHROMATIN.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P26686-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P26686-2; Sequence=VSP_005878;
CC -1- DEVELOPMENTAL STAGE: Expressed throughout development.
CC -1- PTM: Extensively phosphorylated on serine residues in the RS
CC domain.
CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -1- SIMILARITY: Belongs to the SR family of splicing factors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X58720; CAA41556.1; -.
CC PIR, A40459; A40459.
CC PIR, A40459; CAA44483.1; -.
CC FLYBASE: FBgn004587; B52.
CC GO; GO:001607; C:nuclear speck; IDA.
CC GO; GO:0005634; C:nucleus; IDA.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 2.
CC SMART; SMO0360; RRM; 2.
CC PROSITE; PS50102; RRM; 2.
CC PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
CC Nuclear protein; Phosphorylation; mRNA splicing; RNA-binding;
CC Repeat; Alternative splicing.
CC INIT_MET
CC 0 0
CC DOMAIN 3 73 RNA-BINDING (RRM) 1.
CC DOMAIN 88 96 GLY-RICH (HINGE REGION).
CC DOMAIN 119 192 RNA-BINDING (RRM) 2.
CC DOMAIN 206 356 ARG/SER-RICH (RS DOMAIN).
CC VARSPLIC 318 338 Missing (in isoform short).
CC /FTId=VSP_005878.
CC CONFLICT 74 74 T->S (IN REF. 1).
CC CONFLICT 102 106 MISSING (IN REF. 1).
CC CONFLICT 195 195 A->R (IN REF. 1).
CC CONFLICT 228 228 T->S (IN REF. 1).
CC CONFLICT 260 260 A->R (IN REF. 1).
CC CONFLICT 279 281 APV->RSR (IN REF. 1).
CC CONFLICT 293 293 S->T (IN REF. 1).
CC SEQUENCE 375 AA; 42393 MW; 20BA327CB4A4194E CRC64;

```

Query Match 45.7%; Score 42; DB 1; Length 375;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGRSRGXSGRSQS 19
Db 202 SGRGRSRSSSSRSRS 216

Search completed: September 28, 2004, 06:05:48
Job time : 8.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-4
Perfect score: 92
Sequence: 1 SHQESTRGSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRILUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	97.8	797	4	Q16824
2	90	97.8	990	4	Q15206
3	90	97.8	1218	4	Q05331
4	84	91.3	798	4	Q9H4U3
5	84	91.3	1084	4	Q01212
6	75	81.5	591	4	Q01720
7	75	81.5	687	4	Q9H4U2
8	70	76.1	465	4	Q03838
9	69	75.0	322	4	Q15845
10	50	54.3	800	5	Q15845
11	48	52.2	822	3	Q9P312
12	48	52.2	1229	3	Q86231
13	47	51.1	260	10	Q9FMU3
14	47	51.1	260	10	Q7XD54
15	47	51.1	471	5	Q9UAY0
16	47	51.1	619	13	Q7J397

17	47	51.1	983	11	Q8BJ53	Q8BJ53 mus musculu
18	47	51.1	983	11	Q812D4	Q812D4 mus musculu
19	47	51.1	1014	11	Q8B1T6	Q8B1T6 mus musculu
20	47	51.1	1015	11	Q8B1W4	Q8B1W4 mus musculu
21	47	51.1	1035	11	Q812D5	Q812D5 mus musculu
22	47	51.1	1067	11	Q8CIG7	Q8CIG7 mus musculu
23	47	51.1	1105	11	Q8K102	Q8K102 mus musculu
24	47	51.1	1107	11	Q80X50	Q80X50 mus musculu
25	47	51.1	1112	11	Q8B0U1	Q8B0U1 mus musculu
26	47	51.1	1284	5	Q9V9Y3	Q9V9Y3 drosophila
27	46	50.0	373	10	Q23876	Q23876 oryza sativ
28	46	50.0	688	13	Q42375	Q42375 brachydanto
29	46	50.0	836	10	Q9SKZ9	Q9SKZ9 arabidopsis
30	46	50.0	1087	4	Q9BTU3	Q9BTU3 homo sapien
31	46	50.0	1829	16	Q86808	Q86808 streptomyce
32	45.5	49.5	144	5	Q819K6	Q819K6 drosophila
33	45.5	49.5	433	5	Q26649	Q26649 strongyloce
34	45	48.9	198	16	Q8VKK4	Q8VKK4 mycobacteri
35	45	48.9	476	12	Q80890	Q80890 herpesvirus
36	45	48.9	652	5	Q77168	Q77168 apis mellif
37	45	48.9	1142	13	Q7SZV2	Q7SZV2 xenopus lae
38	45	48.9	1394	10	Q8H8L9	Q8H8L9 oryza sativ
39	45	48.9	1433	10	Q94ZG3	Q94ZG3 oryza sativ
40	44.5	48.4	319	10	Q94A20	Q94A20 arabidopsis
41	44.5	48.4	553	5	Q9U203	Q9U203 caenorhabdi
42	44.5	48.4	659	5	Q9U202	Q9U202 caenorhabdi
43	44	47.8	157	16	Q7VFP1	Q7VFP1 helicobacte
44	44	47.8	188	10	Q8H8P6	Q8H8P6 oryza sativ
45	44	47.8	222	4	Q96B40	Q96B40 homo sapien

ALIGNMENTS

RESULT 1	Q16824	PRELIMINARY;	PRT;	797 AA.
ID	Q16824;			
AC	Q16824;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Profilaggrin (Fragment).			
GN	FLG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91064347; PubMed=2248957;			
RA	Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;			
RT	"Organization, structure, and polymorphisms of the human profilaggrin			
RT	gene [published erratum appears in Biochemistry 1991 Jun			
RT	11;30(23):5814.];			
RT	11;30(23):5814.];			
RI	Biochemistry 29:9432-9440(1990).			
DR	EMBL; M60502; AAA63248.1; -1990).			
DR	GO; GO:0005196; F:structural molecule activity; IRA.			
DR	InterPro; IPR003303; Filaggrin.			
DR	Pfam; PF03516; Filaggrin; 4.			
DR	PRINTS; PR00487; FILAGGRIN.			
FT	NON TER			
SQ	SEQUENCE 797 AA; 85176 MW; 60B6184763BDA858 CRC64;			
Query Match	97.8%;	Score 90;	DB 4;	Length 797;
Best Local Similarity	94.7%;	Pred. No. 1.6e-06;		
Matches 18;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			
Qy	1 SHQESTRGSRGSRGSGS 19			
Db	427 SHQESTRGSRGSRGSGS 445			
RESULT 2				

```

Q15206 ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AA63244.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 990;
Best Local Similarity 94.7%; Pred. No. 2.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTGRSGXSGSGS 19
DB 227 SHQESTGRSGXSGSGS 245

RESULT 3
Q05331 ID Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipkev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
DE -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DE DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
DE TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
DE -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
DE HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
DE 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
DE DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
CC EMBL; M96943; AA36487.1; -.
CC PIR; A48118; A48118.
CC HSSP; P02593; 1COM.
DR GO; GO:0005856; C:Cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KM Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
KM Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 1218;
Best Local Similarity 94.7%; Pred. No. 2.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTGRSGXSGSGS 19
DB 449 SHQESTGRSGXSGSGS 467

RESULT 4
Q094U3 ID Q094U3 PRELIMINARY; PRT; 798 AA.
AC Q094U3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 798
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 1.7e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTGRSGXSGSGS 19
DB 428 SHQESTGRSGXSGSGS 446

RESULT 5
Q01212 ID Q01212 PRELIMINARY; PRT; 1084 AA.

```

```

AC Q01212; Q03840;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published extratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814.];"
RL Biochemistry 29:9432-9440(1990).
DR EMBL: M60503; AAA63243.1; JOINED.
DR EMBL: M60501; AAA63243.1; JOINED.
DR GO: GO:0005882; C:Intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007275; P:development; NAS.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF03516; Filaggrin; 6.
DR PRINTS: PR00487; Filaggrin.
FT NON_TER 1
SQ SEQUENCE 1084 AA; 115271 MW; 800C46408BD5A362D CRC64;

Query Match
Best Local Similarity 91.3%; Score 84; DB 4; Length 1084;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
Db 64 SHOESTRGRSGRSGSGS 82

RESULT 6
Q01720 Q01720 PRELIMINARY; PRT; 591 AA.
ID Q01720;
AC Q01720;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Ntunusukisiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
RT J. Biol. Chem. 267:23772-23781(1992).
RL
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTOOLITICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL: L01089; AAA60177.1; -.
DR EMBL: L01090; AAA60176.1; -.

```

```

DR PIR: A45135; A45135.
DR PIR: A48118; A48118.
DR HSSP: P80511; 1E8A.
DR MIM: 135940; -.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001751; CABP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Filaggrin; 2.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; Filaggrin.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467 POTENTIAL.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FILAGGRIN.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match
Best Local Similarity 81.5%; Score 75; DB 4; Length 591;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
Db 449 SHOESTRGRSGRSGSGS 467

RESULT 7
Q094U2 Q094U2 PRELIMINARY; PRT; 687 AA.
ID Q094U2;
AC Q094U2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL: AL356504; CAC13172.1; -.
DR PIR: A48118; A48118.
DR HSSP: P80511; 1E8A.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR001751; CABP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Filaggrin; 3.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; Filaggrin.
DR PROSITE: PS00018; EF_HAND; 1.
DR PROSITE: PS00303; S100_CABP; 1.
FT NON_TER 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match
Best Local Similarity 81.5%; Score 75; DB 4; Length 687;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 SHOESTRGRSGRSGS 19
 DB 449 SHOESTRGRSGRSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT: 465 AA.

AC 003838.
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE FILAGRIN (PROFILAGRIN) (Fragment).
 GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RE MEDLINE=91064347; PubMed=2248957;

RA Can S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 29:9432-9440(1990).

RM [2]

RP REVISIONS.

RE MEDLINE=91255199; PubMed=2043621;

RA Can S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

CC REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLAVED.

DR EMBL; M60499; AAA63246.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 3.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RE MEDLINE=99101527; PubMed=9886436;

RA Girard-Neuhäuser E., Durieux J.J., Arnaud M., Dalbon P., Sebba M.,

RA Vincent C., Simon M., Sennu T., Masson-Bessiere C.,

RA Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated

RT anti-filaggrin autoantibodies are posttranslationally generated on

RT various sites of (pro)filaggrin by deimination of arginine residues.";

RL J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC2555.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

Query Match
 Best Local Similarity 75.0%; Score 69; DB 4; Length 322;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 10

ID 015845 PRELIMINARY; PRT: 800 AA.

AC 015845.
 DT 01-JAN-1998 (TRENBLREL. 05, Created)
 DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE L1439.4.
 GN L1439.4.

OS Leishmania major.

OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatida; Leishmania.

OX NCBI_TaxID=5664;

RP SEQUENCE FROM N.A.

RC STRAIN=Friledin;

RE MEDLINE=99179887; PubMed=10077609;

RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,

RA Wagness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,

RA Basien P., Fu G., Ivans A., Stuart K.;

RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of

RT protein-coding genes.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).

DR EMBL; AE001274; AAC24675.1; -

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

DR FIR; 102852; 102852.

Query Match
 Best Local Similarity 54.3%; Score 50; DB 5; Length 800;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

RESULT 11

ID 09P312 PRELIMINARY; PRT: 822 AA.

AC 09P312.
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.
 GN B12P1.10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RE MEDLINE=91064347; PubMed=2248957;

RA Can S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 29:9432-9440(1990).

RM [2]

RP REVISIONS.

RE MEDLINE=91255199; PubMed=2043621;

RA Can S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

CC REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLAVED.

DR EMBL; M60499; AAA63246.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 3.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

FT NON_TER 1

OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 RX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL390091; CAB98213.1; -.
 DR PIR; T51049; T51049.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR005120; Smg4_UbF3.
 DR Pfam; PF00076; rrm1_1.
 DR Pfam; PF03467; Smg4_UbF3; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;
 Query Match 52.2%; Score 48; DB 3; Length 822;
 Best Local Similarity 52.9%; Pred. No. 19;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 QESTGRSGXSGRSGS 19
 Db 414 RESASGRTRRRGCGT 430

RESULT 12
 ID 086ZHL PRELIMINARY; PRT; 1229 AA.
 AC 086ZHL;
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Related to cell division cycle 2-related protein kinase 7.
 GN 754.060.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 RX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX294020; CAD70910.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0000910; F:cyclin kinase; IEA.
 DR InterPro; IPR000719; Prc1_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kin_AS.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Pkinase_I.
 DR Pfam; PF000001; Prc1_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TYKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW Cell division; Kinase.
 SQ SEQUENCE 1229 AA; 136776 MW; 43BE61265EF73E4 CRC64;
 Query Match 52.2%; Score 48; DB 3; Length 1229;
 Best Local Similarity 57.9%; Pred. No. 29;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 SHQESTGRSGXSGRSGS 19
 Db 242 SHHERRRDRKRRSGRSGS 260

RESULT 13
 ID 09FWU3 PRELIMINARY; PRT; 260 AA.
 AC 09FWU3;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Putative secretory protein.
 GN OSUNBA0051D19.12.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 RX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Varaklen S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldtlyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSUNBA0051D19 genomic sequence."
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC023240; AAG33529.1; -.
 DR Gramene; Q9FWU3; -.
 DR InterPro; IPR007541; BSP.
 DR Pfam; PF04450; BSP; 1.
 SQ SEQUENCE 260 AA; 28706 MW; CDC696A4EC58B022 CRC64;
 Query Match 51.1%; Score 47; DB 10; Length 260;
 Best Local Similarity 69.2%; Pred. No. 7.9;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 6 TRGRSRGSGRSG 18
 Db 35 SKGRSRGAHGRRG 47

RESULT 14
 ID 07XD54 PRELIMINARY; PRT; 260 AA.
 AC 07XD54;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Putative secretory protein.
 GN OSUNBA0051D19.12.
 OS Oryza sativa (Japanese cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 RX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569 (2003).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Ming R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE017108; AAP54393.1; -
 SQ SEQUENCE 260 AA; 28706 MW; CDC696A4EC58B022 CRC64;

Query Match 51.1%; Score 47; DB 10; Length 260;
 Best Local Similarity 69.2%; Pred. No. 7.9;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 TRGSRGSGSGSG 18
 : |||||
 Db 35 SRGSRGAGRRG 47

RESULT 15

Q9UAY0 PRELIMINARY; PRT; 471 AA.
 AC Q9UAY0;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE W03G1.5 protein.
 GN W03G1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditie.
 CX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Riken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A., Scheet P., Harper M.;
 RT "The sequence of C. elegans cosmid W03G1."
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF125964; AAD14753.1; -
 DR PIR: T33987; T33987.
 DR WormPep; W03G1.5; CE17283.
 SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A985B CRC64;

Query Match 51.1%; Score 47; DB 5; Length 471;
 Best Local Similarity 52.6%; Pred. No. 15;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTRGSRGSGSGSGS 19
 : |||||
 Db 167 SSRSPSRGRGSGSGSGS 185

Search completed: September 28, 2004, 06:12:42
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-4
Perfect score: 92
Sequence: 1 SHQESTRGSRGSGRSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_29Jan04:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	2	AAW61508 Peptide c
2	90	97.8	19	2	AAW61514 Peptide c
3	90	97.8	1467	5	ABB97605 Novel hum
4	85	93.5	19	2	AAW61517 Peptide c
5	85	92.4	19	2	AAW61516 Peptide c
6	84	91.3	19	2	AAW61506 Peptide c
7	84	91.3	19	2	AAW61515 Peptide c
8	84	91.3	19	2	AAW61507 Peptide c
9	84	91.3	19	2	AAW61512 Peptide c
10	84	91.3	19	2	AAW61509 Peptide c
11	84	91.3	19	2	AAW61505 Peptide c
12	78	84.8	19	2	AAW61511 Peptide c
13	78	84.8	19	2	AAW61513 Peptide c
14	78	84.8	19	2	AAW61510 Peptide c
15	75	81.5	21	2	AAW61520 Peptide c
16	75	81.5	22	4	AAE07235 IGP1546 P
17	66	71.7	330	2	AAV22955 Human fil
18	66	71.7	330	2	AAV22956 Human fil
19	66	71.7	330	2	AAV22957 Human fil
20	66	67.4	330	2	AAV22954 Human fil
21	56	60.9	477	6	ABO07142 Novel hum
22	51	55.4	1711	4	AAW79819 Human pro
23	51	55.4	1951	4	AAW78835 Human pro
24	49	53.3	641	4	ABG19110 Novel hum
25	48	52.2	772	4	AA664573 Human Mre

26	47	51.1	184	3	AAW57041 Human pro
27	47	51.1	1199	4	ABB58274 Drosophi
28	46	50.0	134	4	ABG23065 Novel hum
29	46	50.0	318	6	ABU70492 Human adi
30	46	50.0	903	6	ABU70701 Human adi
31	45.5	49.5	135	4	ABB66181 Drosophi
32	44	47.8	124	4	ABB72046 Drosophi
33	44	47.8	222	4	AAW93652 Human pol
34	44	47.8	229	7	ADE59825 Rat Prote
35	44	47.8	229	7	ADE59833 Rat Prote
36	44	47.8	229	7	ADE59837 Rat Prote
37	44	47.8	229	7	ADE59829 Rat Prote
38	44	47.8	285	4	ABB69359 Drosophi
39	44	47.8	826	3	AAV96513 Human Zsi
40	44	47.8	837	4	ABG19458 Novel hum
41	44	47.8	1038	7	ADCC03412 Rice flow
42	44	47.8	1308	6	ABR57072 Human tau
43	44	47.8	1331	7	ADCI0010 Human NOV
44	43	46.7	173	6	ABR41796 Human DIT
45	43	46.7	356	5	AAE15535 Beta vulg

ALIGNMENTS

RESULT 1	AAW61508	standard; peptide, 19 AA.
ID	AAW61508	standard; peptide, 19 AA.
XX	AAW61508;	
AC	AAW61508;	
DT	26-OCT-1998	(first entry)
XX	Peptide cfc4, based on cDNA of a profilaggrin repeat.	
DE	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
XX	solid phase synthesis; peptide amide; polyclonal antibody;	
KW	monoclonal antibody.	
XX	Synthetic.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	Modified-site	13
FT	FT	/note= "Citruiline"
XX	WO9822503-A2.	
XX	28-MAY-1998.	
PD	14-NOV-1997;	97WO-NL000624.
XX	15-NOV-1996;	96NL-01004539.
XX	(SCHE-) SITTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
PA	(TEWE-) SITTING TECH WETENSCHAPPEN.	
XX	Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;	
PI	WPI, 1998-398613/34.	
XX	Peptide derived from an antigen recognised by autoantibodies - is	
DR	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
PT	used in diagnosis of the disease.	
XX	Disclosure; Page 6; 19pp; English.	
PS	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of	
XX	the profilaggrin antigen which is recognised by autoantibodies from	
CC	patients with Rheumatoid arthritis (RA). This peptide is reactive with a	
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	The peptides were created by using standard solid phase synthesis, which	
CC	produced them as peptide amides. These sequences may be used in the	

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 1 SHQESTRGRSGXSGRSGS 19

RESULT 2

AAW61514
 ID AAW61514 standard; peptide, 19 AA.

AC AAW61514;

DT 26-OCT-1998 (first entry)

DE Peptide cf, based on cDNA of a proflaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 2.1e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 1 SHQESTRGRSGXSGRSGS 19

RESULT 3

ABB97605
 ID ABB97605 standard; protein; 1467 AA.

XX ABB97605;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

KW Human; anti-neoplastic; vulnery; anti-inflammatory; immunomodulator;
 KW anti-infectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RR;

DR WPI; 2002-292408/33.

DR N-PSDB; ABB97605.

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.

XX Example 2; SEQ ID NO 873; 509pp; English.

CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX Sequence 1467 AA;

Query Match 97.8%; Score 90; DB 5; Length 1467;
 Best Local Similarity 94.7%; Pred. No. 1.5e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGXSGRSGS 19
 |||||
 DB 449 SHQESTRGRSGXSGRSGS 467

RESULT 4
 AAW61517
 ID AAW61517 standard; peptide, 19 AA.

XX AAW61517;

DT 26-OCT-1998 (first entry)

DE Peptide cf, based on cDNA of a proflaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.


```

OS Synthetic.
OS Homo sapiens.
XX WO9822503-A2.
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 93.5%; Score 86; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 9.1e-07;
XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 SHOESTGRSRGRSGRSGS 19
   |||||:|||||:|||||
Db 1 SHOESTGRSRGRSGRSGS 19

```

```

XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 92.4%; Score 85; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.3e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 SHOESTGRSRGRSGRSGS 19
   |||||:|||||:|||||
Db 1 SHOESTGRSRGRSGRSGS 19

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 7
 AAW61515
 ID AAW61515 standard; peptide; 19 AA.

AC AAW61515;

DT 26-OCT-1998 (first entry)

DE Peptide cFA, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

OS Synthetic.

OS Homo sapiens.

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

Db 1 SHOESTRGRSGRSGSGS 19

RESULT 8

AAW61507
 ID AAW61507 standard; peptide; 19 AA.

AC AAW61507;

DT 26-OCT-1998 (first entry)

DE Peptide cfc3, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

OS Synthetic.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Modified-site 11 /note= "Citrulline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
 Db 1 SHOESTRGRSGRSGSGS 19

RESULT 9
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 AC AAW61512;
 DT 26-OCT-1998 (first entry)

```

DE Peptide cfc6, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 7
FT Modified-site /note= "Citruilline"
FT Modified-site 13
FT Modified-site /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 1.9e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SHOESTGRSRGSGSGS 19
DB 1 SHOESTGRSRGSGSGS 19
RESULT 10
AAM61509
ID AAM61509 standard; peptide; 19 AA.
XX
XX AAM61509;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc5, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT Modified-site 16
FT Modified-site /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTGRSRGSGSGS 19
DB 1 SHOESTGRSRGSGSGS 19
RESULT 11
AAM61505
ID AAM61505 standard; peptide; 19 AA.
XX
XX AAM61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Modified-site 7
FT Modified-site /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX

```


CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX
 XX
 SQ Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRGRSGRSGS 19
 |||||
 Db 1 SHOESTXGRSGRSGSGS 19

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 DT 26-OCT-1998 (first entry)
 DE Peptide cfc6, based on cDNA of a proflilagrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97MO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.7e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGRSGS 19
 |||||
 Db 1 SHOESTXGRSGRSGRSGS 19

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 DT 26-OCT-1998 (first entry)
 DE Peptide XI based on cDNA of a proflilagrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruiline"

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97MO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the proflilagrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 5.6e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HOESTGRSGRSGRSGS 19
 |||||
 Db 4 HOESTXGRSGRSGRSGS 21

Search completed: September 28, 2004, 06:24:13
 Job time : 55.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds

(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-4

Percent score: 92

Sequence: 1 SHQESTRGSRGSGSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-4
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-3
8	84	91.3	19	9	US-09-308-150-5
9	84	91.3	19	9	US-09-308-150-8
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-7
13	78	84.8	19	9	US-09-308-150-9
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	56	60.9	477	15	US-10-161-927-62	Sequence 62, Appl
17	51	55.4	78	16	US-10-437-963-203073	Sequence 203073,
18	50	54.3	133	16	US-10-437-963-136367	Sequence 136367,
19	48	53.3	189	16	US-10-437-963-181020	Sequence 181020,
20	48	52.2	772	15	US-10-369-493-6474	Sequence 6474, Ap
21	47	51.1	123	16	US-10-767-701-3436	Sequence 3436, A
22	47	51.1	156	16	US-10-437-963-200529	Sequence 200529,
23	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
24	47	51.1	258	16	US-10-437-963-140693	Sequence 140693,
25	47	51.1	952	16	US-10-437-963-173648	Sequence 173648,
26	46.5	50.5	294	12	US-10-425-114-48469	Sequence 48469, A
27	46.5	50.5	332	12	US-10-425-114-66862	Sequence 66862, A
28	46	50.0	373	16	US-10-437-963-143403	Sequence 143403,
29	46	50.0	466	15	US-10-369-493-8178	Sequence 8178, Ap
30	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
31	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
32	45	48.9	49	12	US-10-424-599-187406	Sequence 187406,
33	45	48.9	105	16	US-10-437-963-196208	Sequence 196208,
34	45	48.9	128	16	US-10-437-963-111350	Sequence 111390,
35	45	48.9	162	16	US-10-437-963-137223	Sequence 137223,
36	45	48.9	248	16	US-10-437-963-185716	Sequence 185716,
37	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A
38	45	48.9	1232	16	US-10-437-963-192222	Sequence 192222,
39	45	48.9	1332	15	US-10-374-780A-1037	Sequence 1037, Ap
40	45	48.9	1380	16	US-10-437-963-185720	Sequence 185720,
41	45	48.9	1394	16	US-10-437-963-185722	Sequence 185722,
42	45	48.9	1409	15	US-10-374-780A-1038	Sequence 1038, Ap
43	45	48.9	1409	16	US-10-437-963-110684	Sequence 110684,
44	45	48.9	1433	15	US-10-374-780A-1035	Sequence 1035, Ap
45	45	48.9	1433	15	US-10-374-780A-1040	Sequence 1040, Ap

ALIGNMENTS

RESULT 1
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoel, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 2

US-09-308-150-11
; Sequence 11, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 2,6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 3

US-09-308-150-14
; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 9,9e-06;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 4

US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1,4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSRGSGSGS 19
Db 1 SHOESTRGRSRGSGSGS 19

RESULT 5

US-09-308-150-1
; Sequence 1, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM


```
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 6
US-09-308-150-2
/ Sequence 2, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 7
US-09-308-150-3
```

```
/ Sequence 3, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3
```

```
Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

```
RESULT 8
US-09-308-150-5
/ Sequence 5, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Scheellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoet, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ CURRENT FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: Known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
```

```
Qy      1 SHOESTRGSRGSRGSGS 19
Db      1 SHOESTRGSRGSRGSGS 19
```

US-09-308-150-5

Query Match 91.3%; Score 84; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTGRSRGSGSGS 19
 Db 1 SHOESTGRSRGSGSGS 19

RESULT 9

US-09-308-150-8
 Sequence 8, Application US/09308150
 Patent No. US20020137092A1

GENERAL INFORMATION:
 APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: NL 1004539
 PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 8
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Derived from
 OTHER INFORMATION: Known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-8

Query Match 91.3%; Score 84; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 1.9e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTGRSRGSGSGS 19
 Db 1 SHOESTGRSRGSGSGS 19

RESULT 10

US-09-308-150-12
 Sequence 12, Application US/09308150
 Patent No. US20020137092A1
 GENERAL INFORMATION:
 APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 12
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Derived from
 OTHER INFORMATION: Known cDNA sequences of human profilaggrin
 US-09-308-150-12

Query Match 91.3%; Score 84; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.9e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTGRSRGSGSGS 19
 Db 1 SHOESTGRSRGSGSGS 19

RESULT 11

US-09-308-150-6
 Sequence 6, Application US/09308150
 Patent No. US20020137092A1
 GENERAL INFORMATION:
 APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius
 APPLICANT: Raats, Jozef Maria Hendrik
 APPLICANT: Hoeft, Rene Michael Antonius
 APPLICANT: Stichting Scheikundig Onderzoek Nederland
 TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 CURRENT APPLICATION NUMBER: US/09/308,150
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: PCT/NL97/00624
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: NL 1004539
 PRIOR FILING DATE: 1996-11-15
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 6
 LENGTH: 19
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Derived from
 OTHER INFORMATION: Known cDNA sequences of human profilaggrin
 OTHER INFORMATION: Xaa is citrulline
 US-09-308-150-6

Query Match 84.8%; Score 78; DB 9; Length 19;
 Best Local Similarity 84.2%; Pred. No. 0.00014;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSRGSGSGS 19
 Db 1 SHOESTGRSRGSGSGS 19

RESULT 12

US-09-308-150-7
 Sequence 7, Application US/09308150
 Patent No. US20020137092A1
 GENERAL INFORMATION:
 APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSGRSGS 19
Db 1 SHOESTGSRGSRGSGRSGS 19

RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: AUTOANTIBODIES DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTRGSRGSGRSGS 19
Db 1 SHOESTGSRGSRGSGRSGS 19

RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
NAME/KEY: DISULFID
LOCATION: (3)..(16)
US-09-308-150-10

Query Match 81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00043;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HOESTRGSRGSGRSGS 19
Db 4 HOESTGSRGSRGSGRSGS 21

RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydia
TITLE OF INVENTION: PEPTIDS DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.003INFUS00 INNS:031
CURRENT FILING DATE: 2000-12-21
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US/09/747,029A
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptides
; NAME/KEY: MOD_RES
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22
```

```
Query Match      81.5%; Score 75; DB 9; Length 22;
Best Local Similarity 83.3%; Pred. No. 0.00045;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      2 HOESTRGRSRGSGRSGS 19
          ||||| ||||| |||||
Db       5 HOESTYGRSRGRCGRSGS 22
```

Search completed: September 28, 2004, 07:28:50
Job time : 111.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQESTRGRSGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Dirl:*
2: Dirl:*
3: Dirl:*
4: Dirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	86	93.5	2248	2 A35938	profilaggrin - hum
2	84	91.3	416	2 A32947	filaggrin precursor
3	75	81.5	591	2 A45135	profilaggrin - hum
4	58	63.0	506	1 W2ML47	E2 protein - human
5	48	52.2	772	2 T27512	hypothetical prote
6	48	52.2	822	2 T51049	related to nucleol
7	47	51.1	308	2 A29379	glycine-rich prote
8	47	51.1	373	2 T02976	probable DNA bindi
9	47	51.1	471	2 T33997	hypothetical prote
10	46	50.0	467	2 T01462	hypothetical prote
11	46	50.0	759	2 F86362	hypothetical prote
12	45	50.0	836	2 G84727	probable DNA topoi
13	45	48.9	409	2 T35118	probable secreted
14	45	48.9	1703	2 S15047	SNF2 protein - yea
15	45	48.4	682	2 T15092	hypothetical prote
16	44.5	48.4	687	1 S13654	ATP-dependent RNA
17	44.5	48.4	1496	2 T00499	probable retroelem
18	44	47.8	229	2 JC7219	nuclear protein SR
19	44	47.8	452	2 S36482	E2 protein - human
20	44	47.8	483	2 S36470	E2 protein - human
21	44	47.8	485	2 A40751	finger protein MZF
22	44	47.8	867	2 T27136	hypothetical prote
23	44	47.8	871	2 T27135	hypothetical prote
24	44	47.8	1232	2 S40766	hypothetical prote
25	43.5	47.3	463	2 T51194	hypothetical prote
26	43	46.7	123	2 T16234	hypothetical prote
27	43	46.7	151	2 T21827	hypothetical prote
28	43	46.7	210	2 C84404	hypothetical prote
29	43	46.7	237	2 T13649	hypothetical prote

30	43	46.7	240	2 A46179	U2 snRNP auxiliary
31	43	46.7	488	2 S36570	E2 protein - human
32	43	46.7	514	1 W2ML5	E2 protein - human
33	43	46.7	695	2 T51652	desRNA-binding prot
34	43	46.7	718	2 T02531	probable SCARECROW
35	43	46.7	745	2 G01025	serine/threonine p
36	43	46.7	774	2 T48609	probable serine/th
37	43	46.7	776	2 A87330	sensory box histid
38	43	46.7	800	2 T02852	probable membrane
39	43	46.7	800	2 T51653	desRNA-binding prot
40	43	46.7	843	2 H86209	protein F22G5.10 l
41	43	46.7	1119	2 T16720	hypothetical prote
42	43	46.7	1339	2 T47841	hypothetical prote
43	43	46.7	2233	2 S63347	acetyl-CoA carboxy
44	42	45.7	166	2 T39586	rna binding protei
45	42	45.7	184	2 S26046	sex-determining pr

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gam, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; MUID:91064347; PMID:2248957
A/Accession: A35938
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:J02929
C/Genetics:
A/Genes: GDB:FIG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:246-569/Region: profilaggrin repeat
F:570-893/Region: profilaggrin repeat
F:1074-1197/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 93.5%; Score 86; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 4.3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGRSGRSGXSGS 19
DB 551 SHQESTRGRSGRSGXSGS 569

RESULT 2
A32947
filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinnley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of t
A/Reference number: A32947; MUID:89296601; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCR>
A/Cross-references: GB:M2435; NID:9182604; PIDN:AA52454.1; PID:9182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 91.3%; Score 84; DB 2; Length 416;
Best Local Similarity 89.5%; Pred. No. 1.7e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19
DB 7 SHOESTRGRSGRSGSGS 25

RESULT 3
A45135
Profilaggrin - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A45135
R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nurnsukstiri, W.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992
A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A:Reference number: A45135; MUID:93054736; PMID:1429717
A:Accession: A45135
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-591 <PRE>
A:Cross-references: GB:101089; GB:M90967; NID:9190408; PIDN:AAA60177.1; PID:553621
A>Note: sequence extracted from NCBI backbone (NCBIF:118773)
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 81.5%; Score 75; DB 2; Length 591;
Best Local Similarity 84.2%; Pred. No. 0.0006;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19
DB 449 SHOESTRGRSGRSGSGS 467

RESULT 4
M2ML47
E2 protein - human papillomavirus type 47
C:Species: human papillomavirus type 47
A>Note: host Homo sapiens (man)
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: D35324
R:Kiyono, T.; Adachi, A.; Ishibashi, M.
Virology 177, 401-405, 1990
A:Title: Genome organization and taxonomic position of human papillomavirus type 47
A:Reference number: A35324; MUID:90281611; PMID:2162112
A:Accession: D35324
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-506 <KTY>
A:Cross-references: GB:M42305; NID:9333062; PIDN:AAA46979.1; PID:9333067
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 63.0%; Score 58; DB 1; Length 506;
Best Local Similarity 57.9%; Pred. No. 0.22;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19

DB 342 SRENTTRGRGRGRGAGS 360

RESULT 5
T27512
Hypothetical protein ZC302.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2003
C:Accession: T27512
R:Kelly, P.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z20380
A:Accession: T27512
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-772 <MTL>
A:Cross-references: EMBL:Z73978; PIDN:CAA98292.1; GSPDB:GN00023; CESP:ZC302.1
A:Experimental source: clone ZC302
C:Genetics:
A:Gene: CESP:ZC302.1
A:Map position: 5
A:Introns: 43/2; 70/3; 94/2; 129/2; 393/3; 548/3; 598/3; 642/3
C:Superfamily: double-strand break repair protein MRE11; phosphodiesterase core homology

Query Match 52.2%; Score 48; DB 2; Length 772;
Best Local Similarity 52.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

CY 1 SHOESTRGRSGRSGSGS 19
DB 697 SKQPTTRGRGRGAGAGS 715

RESULT 6
T51049
related to nucleolar phosphoprotein [imported] - Neurospora crassa
N:Alternate names: protein B12Fl.10
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51049
R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51049
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-822 <SCH>
A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12Fl.10
A:Experimental source: BAC clone B12Fl; strain OR74A
C:Genetics:
A:Gene: NCSP:B12Fl.10
A:Map position: 6
A:Introns: 80/2

Query Match 52.2%; Score 48; DB 2; Length 822;
Best Local Similarity 52.9%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 3 OESTRGRSGRSGSGS 19
DB 414 RESASGRTRGRGRGT 430

RESULT 7
A29379
glycine-rich protein GRP3 - brine shrimp
C:Species: Artemia salina (brine shrimp)
C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Feb-1997
C:Accession: A29379
R:Cruz-Alvarez, M.; Pellicer, A.
J. Biol. Chem. 262, 13377-13380, 1987
A:Title: Cloning of a full-length complementary DNA for an Artemia salina glycine-rich p
A:Reference number: A29379; MUID:88007550; PMID:2443491

A:Accession: A29379
A:Molecule type: mRNA
A:Residues: 1-308 <CRU>
C:Keywords: nucleus

Query Match 51.1%; Score 47; DB 2; Length 308;
Best Local Similarity 75.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 RGRSGRSGSGS 18
Db 210 RGRSGRSGSGS 221

RESULT 8

T02976
probable DNA binding protein PCF2 - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000

C:Accession: T02976
R:Kosugi, S.; Ohashi, Y.
Plant Cell 9, 1607-1619, 1997
A>Title: PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell
A:Reference number: Z14803; MUID:97480096; PMID:9338963
A:Accession: T02976
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-373 <ROS>
A:Cross-references: EMBL:DB7261; NID:G2580439; PIDN:BA21143.1; PID:G2580440
A:Experimental source: cultivar Nipponbare

Query Match 51.1%; Score 47; DB 2; Length 373;
Best Local Similarity 53.3%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGS 15
Db 334 SHEQGRGRSGRSGS 348

RESULT 9

T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T33997
R:Pauley, A.; Schneet, P.; Harper, M.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-471 <PNU>
A:Cross-references: EMBL:AF129964; PIDN:AA14753.1; GSPDB:GN00022; CESP:W03G1.5
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 51.1%; Score 47; DB 2; Length 471;
Best Local Similarity 52.6%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHOESTRGRSGRSGSGS 19
Db 167 SRSRSGRSGRSGSGS 185

RESULT 10
T01462
hypothetical protein T24H24.19 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01462
R:Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K.
submitted to the EMBL Data Library, August 1998
A:Description: The sequence of A. thaliana T24H24.
A:Reference number: Z14333
A:Accession: T01462
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-467 <COU>
A:Cross-references: EMBL:AF075598; NID:G3293581; PID:G3377836
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A>Note: T24H24.19

Query Match 50.0%; Score 46; DB 2; Length 467;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 2 HOESTRGRSGRSGSGS 19
Db 188 HNNRGRGRSGRSGSGS 205

RESULT 11

F86362
F19G10.17 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001

C:Accession: F86362
R:Rheologos, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, C.M.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86362
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-759 <STO>
A:Cross-references: GB:AE005172; NID:G2462837; PIDN:AB72172.1; GSPDB:GN00141
A:Map position: 1

Query Match 50.0%; Score 46; DB 2; Length 759;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 RGRSGRSGSGS 17
Db 728 RGRSGRSGSGS 738

RESULT 12

G84727
probable DNA topoisomerase III beta [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84727
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; P. eus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. eus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84727
A>Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-836 <STD>
 A:Cross-references: GB:AE02093; NID:G4263718; PIDN:AAD15404.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2932000
 A:Map position: 2

Query Match 50.0%; Score 46; DB 2; Length 836;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 RGRSRGRSGXSGS 19
 ||| ||| ||| |||
 Db 812 RGRSRGRSGXSGS 824

RESULT 13

T15118
 probable secreted protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 08-Sep-2000
 C:Accession: T15118
 R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, March 1998
 A:Reference number: 221568
 A:Accession: T15118
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-409 <SEP>
 A:Cross-references: EMBL:AL022268; PIDN:CAA18335.1; GSPDB:GN00070; SCOEEDB:SC4H2.19C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEEDB:SC4H2.19C
 C:Superfamily: Streptomyces coelicolor probable secreted protein SC4H2.19C

Query Match 48.9%; Score 45; DB 2; Length 409;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 5 STRGRSRGRSGXSGS 19
 ||| ||| ||| |||
 Db 9 STRGRSRGRSGXSGS 23

RESULT 14

S15047
 SNF2 protein - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein O5509; protein O5509c; protein YOR290C
 C:Species: Saccharomyces cerevisiae
 C>Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Sep-1999
 C:Accession: S15047; S16820; S67192; S67194; S72058
 R:Laurent, B.C.; Treitel, M.A.; Carlson, M.
 Proc. Natl. Acad. Sci. U.S.A. 88, 2687-2691, 1991
 A:Title: Functional interdependence of the yeast SNF2, SNF5, and SNF6 proteins in trans
 A:Reference number: S15047; MUID:91187857; PMID:1901413
 A:Accession: S15047

A:Molecule type: DNA
 A:Residues: 1-1703 <PRO>
 A:Cross-references: EMBL:M61703; NID:G172631; PIDN:AA35059.1; PID:G172632
 R:Yoshimoto, H.; Yamashita, I.
 Mol. Gen. Genet. 228, 270-280, 1991
 A:Title: The GML/SNF2 gene of Saccharomyces cerevisiae encodes a highly charged nuclear
 A:Reference number: S16820; MUID:91360076; PMID:1886612
 A:Accession: S16820
 A:Molecule type: DNA
 A:Residues: 1-1703 <YOS>
 A:Cross-references: EMBL:X57837; NID:G4499; PIDN:CAA40969.1; PID:G4500
 R:Cheret, G.; Sor, F.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67169
 A:Accession: S67169
 A:Molecule type: DNA
 A:Residues: 1-1703 <CHE>

A:Cross-references: EMBL:Z75198; NID:G1420643; PIDN:CAA9517.1; PID:e252424; PID:G1420644
 A:Experimental source: strain S288C
 R:Utzlich, C.; Jannaux, J.C.; Kordas, E.; Poirey, R.; Pujol, A.; Tobiasch, E.
 Submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67194
 A:Accession: S67194

A:Molecule type: DNA
 A:Residues: 1-108 <CZI>
 A:Cross-references: EMBL:Z75198; MIPS:YOR290C
 A:Experimental source: strain S288C
 R:Cheret, G.; Bernardi, A.; Sor, F.
 Yeast 12, 1059-1064, 1996
 A:Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces
 A:Reference number: S72039; MUID:97051594; PMID:8896271
 A:Accession: S72039
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-1703 <CHW>
 A:Cross-references: EMBL:X89633; NID:G1279694; PIDN:CAA61793.1; PID:e189413; PID:G127971
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
 C:Keywords: nucleus; transcription regulation
 C:Keywords: nucleus; transcription regulation
 F:1576-1631/Domain: bromodomain homology <BRO>

Query Match 48.9%; Score 45; DB 2; Length 1703;
 Best Local Similarity 55.6%; Pred. No. 71;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHQESTRGRSRGRSGXSG 18
 ||| ||| ||| |||
 Db 1499 SARTSTRGRGRGRGRG 1516

RESULT 15

T15092
 hypothetical protein K06A5.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15092
 R:Wamsley, P.
 submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid K06A5.
 A:Reference number: Z18291
 A:Accession: T15092

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-662 <WAM>
 A:Cross-references: EMBL:AF039038; NID:G2736359; PID:G2736362; PIDN:AAB94170.1; GSPDB:GNT
 A:Experimental source: strain Bristol N2; clone K06A5
 C:Genetics:
 A:Gene: CESP:K06A5.1
 A:Map position: 1
 A:Introns: 5/3; 47/3; 157/3; 509/3; 590/3; 646/3

Query Match 48.4%; Score 44.5; DB 2; Length 682;
 Best Local Similarity 61.1%; Pred. No. 35;
 Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy 5 STRGRSRGRSGXSGS 19
 ||| ||| ||| |||
 Db 358 TTRSRGRGRFELSGNSGA 375

Search completed: September 28, 2004, 06:15:00
 Job time : 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 : Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQSTRGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	91.3	416	1	FILA_HUMAN
2	58	63.0	506	1	VE2_HPV47
3	48	52.2	772	1	MR11_CAEEL
4	47	51.1	308	1	GRP3_ARTSA
5	45	48.9	70	1	L2W0_ADE40
6	45	48.9	1703	1	SNF2_YEAST
7	44.5	48.4	697	1	AN3_XENLA
8	44	47.8	452	1	VE2_HPV17
9	44	47.8	483	1	VE2_HPV14
10	44	47.8	734	1	ZN42_HUMAN
11	44	47.8	1232	1	Y005_CAEEL
12	43	46.7	123	1	LSM4_CAEEL
13	43	46.7	210	1	Y043_HALNI
14	43	46.7	239	1	U2AG_MOUSE
15	43	46.7	240	1	U2AG_MOUSE
16	43	46.7	488	1	VE2_HPV49
17	43	46.7	514	1	VE2_HPV05
18	43	46.7	774	1	MRK2_MOUSE
19	43	46.7	2233	1	COAC_YEAST
20	42	45.7	184	1	TRSF_DROSI
21	42	45.7	197	1	TRSF_DROME
22	42	45.7	694	1	FRZ2_DROME
23	42	45.7	706	1	MR11_MOUSE
24	42	45.7	1015	1	PRQ_CRASP
25	42	45.7	2404	1	SON_MOUSE
26	42	45.7	2426	1	SON_HUMAN
27	41.5	45.1	1180	1	TYK2_MOUSE
28	41	44.6	176	1	SSB_TREPA
29	41	44.6	197	1	RX21_DROME
30	41	44.6	315	1	SOL2_YEAST
31	41	44.6	325	1	SOL2_MOUSE
32	41	44.6	332	1	Z265_RAT
33	41	44.6	337	1	Z265_HUMAN

34	41	44.6	503	1	VE2_HPV21	P50767 human papill
35	41	44.6	525	1	MR11_RAT	P48965 rattus norv
36	41	44.6	526	1	2NBA_YEAST	O00362 saccharomyc
37	41	44.6	655	1	ILF1_HUMAN	Q01167 homo sapien
38	41	44.6	706	1	MR11_RAT	Q91160 rattus norv
39	41	44.6	708	1	MR11_HUMAN	P49955 homo sapien
40	41	44.6	825	1	5S5_RAT	O63003 rattus norv
41	41	44.6	1061	1	PRD4_HUMAN	O94966 homo sapien
42	41	44.6	1337	1	DEXT_STRDO	P39653 streptococc
43	40	43.5	148	1	LSM4_FAGSY	O92809 fagus sylvia
44	40	43.5	206	1	BTJ3_HUMAN	P20290 homo sapien
45	40	43.5	255	1	THO4_MOUSE	O08583 mus musculu

ALIGNMENTS

```

RESULT 1
FILA_HUMAN          STANDARD;          PRT;          416 AA.
ID P20930;
AC 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89296901; PubMed=2740331;
RA McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RA "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4846-4852(1989).
RN
RN [2]
RP CITRULLINATION.
RX MEDLINE=9637438; PubMed=8780679;
RA Senoh T., Kan S., Ogawa H., Manabe M., Asaga H.;
RA "Preferential delamination of keratin K1 and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -!- FUNCTION: Aggregates Keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -!- PTM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -!- PTM: Undergoes delamination of some arginine residues
CC (citrullination).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC
CC -----
CC EMBL: M24355; AAA52454.1; -.
CC PIR: A32947; A32947.
CC Genew: HGNC:3748; FLG.
CC MIM: 135940; -.
CC GO: GO:0005882; C:intermediate filament; NMS.
CC GO: GO:0005198; F:structural molecule activity; NMS.
CC GO: GO:0007275; P:development; NMS.
CC InterPro: IPR003503; Filaggrin.

```

DR Pfam; PF03516; Flaggrin; 2.
 DR PRINTS; PR00487; FLAGGRIN.
 KW Phosphorylation; Citrullination; Developmental protein.
 FT NON TER 1
 SO SEQUENCE 416 AA; 44105 MW; DEEA218BA043F32 CRC64;

Query Match 91.3%; Score 84; DB 1; Length 416;
 Best Local Similarity 89.5%; Pred. No. 6.3e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SH0ESTRGRSGRSGSGS 19
 DB 7 SH0ESTRGRSGRSGSGS 25

RESULT 2

VE2_HPV47 STANDARD; PRT; 506 AA.
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2.
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).
 CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNNCGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2RE'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STRICTLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M32305; AAA46979.1; -.
 DR PIR; D5324; W2ML47.
 DR HSSP; P03122; 2BOP.
 DR InterPro; IPR000427; E2_C.
 DR InterPro; IPR001866; E2_N.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF00511; E2_C; 1.
 DR Pfam; PF00508; E2_N; 1.
 DR ProDom; PD000672; E2_C; 1.
 DR ProDom; PD000678; E2_N; 1.
 KW Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SO SEQUENCE 506 AA; 57478 MW; 92C37F4BF875065 CRC64;

Query Match 63.0%; Score 58; DB 1; Length 506;
 Best Local Similarity 57.9%; Pred. No. 0.089;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 SH0ESTRGRSGRSGSGS 19

DB 342 SREGNTRGRSGRSGSGS 360

RESULT 3
 ID MRE11_CAEEL STANDARD; PRT; 772 AA.

AC Q23255;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Double-strand break repair protein mre-11.
 GN MRE-11 OR ZC302.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
 RP GLUT-139.
 RX MEDLINE=21135651; PubMed=11238374;
 RA Chin G.M., Villeneuve A.M.;
 RT "C. elegans mre-11 is required for meiotic recombination and DNA
 RT repair but is dispensable for the meiotic G(2) DNA damage
 RT checkpoint.";
 RL Genes Dev. 15:522-534(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Kelly P.F.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
 CC processing.
 CC -1- FUNCTION: Required for meiotic crossing over and chiasma
 CC formation. Pachytene morphology and homolog pairing are normal.
 CC Vital in long term for maintenance of reproductive capacity of
 CC subsequent generations.
 CC -1- COPACITOR: Manganese (By similarity).
 CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the MRE11/RAD32 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z73978; CAA98292.1; -.
 DR PIR; T27512; T27512.
 DR GeronOnline; 208612; -.
 DR WormRep; ZC302.1; CE06573.
 DR InterPro; IPR003701; DNA_repair.
 DR InterPro; IPR004843; M-psestrase.
 DR InterPro; IPR007281; Mre11_DNA_bind.
 DR Pfam; PF00149; Metallophos; 1.
 DR Pfam; PF04152; Mre11_DNA_bind; 1.
 DR TIGRfam; TIGR00583; mre11; 1.
 KW DNA repair; Hydrolyase; Nuclease; Endonuclease; Exonuclease;
 KW Nuclear protein; Manganese; Meiosis.
 FT MUTAGEN 139 E->K; IN MRE-11-ME41, DEFECTIVE IN
 FT MEIOTIC CHROMOSOME DEGRADATION.
 SO SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;

Query Match 52.2%; Score 48; DB 1; Length 772;
 Best Local Similarity 52.6%; Pred. No. 5.1;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 SH0ESTRGRSGRSGSGS 19

Db 697 SKOPTTRGRGRGARGAGAS 715

RESULT 4

GRP3_ARTSA STANDARD; PRT; 308 AA.
 ID GRP3_ARTSA STANDARD; PRT; 308 AA.
 AC P13230;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycine-rich protein GRP3.
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 NCBI_TaxID=85549;
 RX MEDLINE=88007550; PubMed=2443491;
 RA Cruz-Alvarez M., Pellicer A.;
 RT "Cloning of a full-length complementary DNA for an Artemia salina
 RT glycine-rich protein. Structural relationship with RNA binding
 RT proteins.";
 RL J. Biol. Chem. 262:13377-13380(1987).
 CC -1- PTM: THE ARGININES IN THE GLY-RICH DOMAIN MIGHT BE METHYLATED.
 CC -1- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 CC HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 CC STRUCTURALLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC -1- SIMILARITY: Contains 1 KH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03453; AAC83400.1; -.
 DR PIR: A29379; A29379.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_Type_1.
 DR SMART: SM00322; KH; 1.
 DR PROSITE: PS50084; KH_Type_1; 1.
 DR Nuclear protein; Ribonucleoprotein; Methylation; RNA-binding.
 FT DOMAIN 83 118 KH.
 FT 186 308 GLY-RICH.
 SQ SEQUENCE 308 AA; 35034 MW; E534CEE552BB3BC2 CRC64;
 Query Match SNF2_YEAST 51.1%; Score 47; DB 1; Length 308;
 Best Local Similarity 75.0%; Pred. No. 2.7;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 5
 L2MU_ADE40 STANDARD; PRT; 70 AA.
 ID L2MU_ADE40 STANDARD; PRT; 70 AA.
 AC Q64858; Q67171;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Late l2 mu core protein precursor (pmu) (Protein X).
 GN px.
 OS Human adenovirus type 40.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 NCBI_TaxID=28284;
 RX MEDLINE=94087748; PubMed=8263936;
 RA Davison A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
 RT "The DNA sequence of adenovirus type 40.";
 RL J. Mol. Biol. 234:1308-1316(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dugan;
 RX MEDLINE=96332521; PubMed=8760431;
 RA Grydsuk J.D., Fortes E., Petric M., Brown M.;
 RT "Common epitope on protein VI of enteric adenoviruses from subgenera
 RT A and F.";
 RL J. Gen. Virol. 77:1811-1819(1996).
 CC -1- FUNCTION: The role of the precursor might be to condense the viral
 CC prochromatin for encapsidation by virtue of the two basic domains
 CC (by similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L19443; AAC13965.1; -.
 DR EMBL: U14651; AAB19000.1; -.
 DR InterPro: IPR008393; Adenovirus_PX.
 DR Pfam: PF05829; Adenovirus_PX; 1.
 DR Core protein; DNA-binding; Late protein.
 FT PROCEP 1 26 BY SIMILARITY.
 FT CHAIN 27 40 LATE L2 MU CORE PROTEIN.
 FT PROCEP 41 70 BY SIMILARITY.
 FT SITE 26 27 CLEAVAGE (BY ADENOVIRUS PROTEASE)
 FT SITE 40 41 CLEAVAGE (BY ADENOVIRUS PROTEASE)
 FT SITE (POTENTIAL).
 FT CONFLICT 31 32 RA -> PP (IN REF. 2).
 FT SEQUENCE 70 AA; 7596 MW; BF621E01518PF69 CRC64;
 Query Match SNF2_YEAST 48.9%; Score 45; DB 1; Length 70;
 Best Local Similarity 66.7%; Pred. No. 1.1;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 6

SNF2_YEAST STANDARD; PRT; 1703 AA.
 ID SNF2_YEAST STANDARD; PRT; 1703 AA.
 AC P22082;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription regulatory protein SNF2 (SWI/SNF complex component SNF2)
 DE (Regulatory protein SWI2) (Regulatory protein GAMI) (Transcription
 DE factor TYS3).
 GN SNF2 OR SWI2 OR GAMI OR TYS3 OR RIC1 OR YOR290C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetaceae;
 OC Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 RX MEDLINE=91187857; PubMed=1901413;
 RA Laurent B.C., Treitel M.A., Carlson M.;
 RT "Functional interdependence of the yeast SNF2, SNF5, and SNF6
 RT proteins in transcriptional activation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2687-2691(1991).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=AH22; PubMed=1886612;
 RX MEDLINE=91360076; PubMed=1886612;
 RA Yoshimoto H., Yamashita I.;
 RT "The GAML/SNF2 gene of *Saccharomyces cerevisiae* encodes a highly
 RT charged nuclear protein required for transcription of the STAI
 RT gene";
 RL Mol. Gen. Genet. 228:270-280(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1B;
 RX MEDLINE=95332261; PubMed=7608126;
 RA Kodaki T., Hosaka K., Nakawa J., Yamashita S.;
 RT "The SNF2/SWI2/GML1/TYS3/RIC1 gene is involved in the coordinate
 RT regulation of phospholipid synthesis in *Saccharomyces cerevisiae*";
 RL J. Biochem. 117:362-368(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97051594; PubMed=8896271;
 RA Cheret G., Bernardi A., Sor F.J.;
 RT "DNA sequence analysis of the VP11-SNF2 region on chromosome XV of
 RT *Saccharomyces cerevisiae*";
 RL Yeast 12:1059-1064(1996).
 RN [5]
 RP SEQUENCE OF 1-309 FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97298310; PubMed=9153758;
 RA Poljey R., Ciepluch C., Tobiasch E., Pujol A., Kordes E.,
 RA Janniaux J.-C.;
 RT "Sequence and analysis of a 36.2 kb fragment from the right arm of
 RT yeast chromosome XV reveals 19 open reading frames including SNF2 (5'
 RT end), CPA1, SLY4, a putative transport ATPase, a putative ribosomal
 RT protein and an SNF2 homologue";
 RL Yeast 13:479-482(1997).
 RN [6]
 RP MUTAGENESIS.
 RX MEDLINE=97025355; PubMed=8871545;
 RA Richmond E., Peterson C.L.;
 RT "Functional analysis of the DNA-stimulated Arpase domain of yeast
 RT SWI2/SNF2";
 RL Nucleic Acids Res. 24:3685-3692(1996).
 CC -1- FUNCTION: Involved in transcriptional activation. The SWI/SNF
 CC complex is required for the induced expression of a large number
 CC of genes. This complex alters chromatin structure to facilitate
 CC binding of gene-specific dedicated transcription factors.
 CC -1- SUBUNIT: Component of the SWI/SNF global transcription activator
 CC complex.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: STRONG, TO DROSOPHILA BRAHMA.
 CC -1- SIMILARITY: Contains 3 A.T hook DNA-binding repeats.
 CC -1- SIMILARITY: Contains 3 A.T hook DNA-binding repeats.
 CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M61703; AAA35059.1; -
 CC EMBL: X57837; CAA40969.1; -
 CC EMBL: D90459; BAA14423.1; -
 CC EMBL: X89633; CAA61793.1; -
 CC EMBL: Z75198; CAA99517.1; -
 CC EMBL: Z75199; CAA99519.1; -
 CC PIR: S15047; S15047;
 CC Germline; 143877; -
 CC TRANSFAC: T02401; -
 CC SCD; S0005815; SNF2;
 CC InterPro; IPR000637; AT_hook.
 CC InterPro; IPR001487; Bromodomain.

DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001550; Helicase_C.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF02178; AT_hook; 2.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR PRINTS; PR00929; ATHOOK.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00384; AT_hook; 2.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; 1.
 DR PROSITE; PS50014; BROMODOMAIN_2; 1.
 KW Transcription regulation; Nuclear protein; Activator; Repeat;
 KW Bromodomain; ATP-binding; Helicase.
 FT DOMAIN 55
 FT NP BIND 207 239
 FT SITE 792 799
 FT NP BIND 894 897
 FT DNA BIND 1446 1456
 FT DNA BIND 1502 1513
 FT DNA BIND 1516 1526
 FT DOMAIN 1568 1638
 FT SEQUENCE 1703 AA; 194050 MW; 84B8C595C8F3E6D CRC64;
 SQ
 Query Match 48.9%; Score 45; DB 1; Length 1703;
 Best Local Similarity 55.6%; Pred. No. 35;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHQSTGRGRSGRSGXSG 18
 Db 1499 SARTSTRGRGRGRGRGRG 1516
 RESULT 7
 AN_XENLA STANDARD; PRT; 697 AA.
 ID AN3_XENLA
 AC P24346;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative ATP-dependent RNA helicase An3.
 GN AN3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 CX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91141586; PubMed=1996140;
 RA Gururajan R., Perry O'Keefe H., Melton D.A., Weeks D.L.;
 RT "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
 RT RNA helicase";
 RL Nature 349:717-719(1991).
 CC -1- FUNCTION: Putative ATP-dependent RNA helicase.
 CC -1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
 CC subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X57328; CAA40605.1; -
 CC PIR: S13654; S13654.
 CC HSP; Q58083; lHVS.
 CC InterPro; IPR001410; DEAD.

```

DR InterPro: IPR000629; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD_1.
DR Pfam: PF00271; Helicase_C_1.
DR SMART: SM00487; DEXDC_1.
DR SMART: SM00490; HELIC_C_1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE_1.
KW Helicase; ATP-binding; RNA-binding.
FT NE_BIND 265 272 ATP (BY SIMILARITY).
FT SITE 388 391 DEAD_BOX.
FT DOMAIN 623 697 GLY/SER-RICH.
SQ SEQUENCE 697 AA; 77302 MW; F3DD23EB60B2E2EF CRC64;

Query Match
Best Local Similarity 52.6%; Score 44.5; DB 1; Length 697;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 2 HOESTRGRSGR-SGXSGS 19
DB 621 HKSSRGRSGRSGSGGA 639

RESULT 8
VE2_HPV17 STANDARD; PRT; 452 AA.
AC P36785;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 17.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10607;
RN (1)
RP MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74469; CAAS2515.1; -.
DR FIR: S36482; S36482.
DR HSSP: P17383; IDHM.
DR InterPro: IPR00427; E2_C.
DR InterPro: IPR001866; E2_N.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF00511; E2_C_1.
DR Pfam: PF00508; E2_N_1.
DR ProDom: PD000672; E2_C_1.
DR ProDom: PD000678; E2_N_1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 452 AA; 51241 MW; 70298AFL6D3E5PS CRC64;

```

```

Query Match
Best Local Similarity 50.0%; Score 44; DB 1; Length 452;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 HOESTRGRSGRSGXSG 18
DB 293 KDSRSPNRGSGSGG 308

RESULT 9
VE2_HPV14 STANDARD; PRT; 483 AA.
AC P36783;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 14.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10605;
RN (1)
RP MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74467; CAAS2503.1; -.
DR FIR: S36470; S36470.
DR HSSP: P17383; IDHM.
DR InterPro: IPR00427; E2_C.
DR InterPro: IPR001866; E2_N.
DR InterPro: IPR009021; Viral_DNA_bd.
DR Pfam: PF00511; E2_C_1.
DR Pfam: PF00508; E2_N_1.
DR ProDom: PD000672; E2_C_1.
DR ProDom: PD000678; E2_N_1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 483 AA; 54938 MW; CF09E23ADF4E3800 CRC64;

Query Match
Best Local Similarity 50.0%; Score 44; DB 1; Length 483;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGXSG 18
DB 327 SSKRSGRSGRSGSGSG 344

RESULT 10
ZN42_HUMAN

```

ID ZN42 HUMAN STANDARD; PRT; 734 AA.
 AC P28698; Q9NR90; Q9DEW2;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Zinc finger protein 42 (Myeloid zinc finger 1) (MZF-1).
 GN ZNF42 OR MZF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RX MEDLINE=91317761; PubMed=1860835;
 RP SEQUENCE FROM N.A. (ISOFORM MZF1A).
 RA Hromas R., Collins S.J., Hickstein D., Raskind W., Deaven L.L.,
 RA O'Hara P., Hagen F.S., Kaushansky K.;
 RT "A retinoic acid-responsive human zinc finger gene, MZF-1,
 RT preferentially expressed in myeloid cells";
 RL J. Biol. Chem. 266:14183-14187 (1991).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM MZF1A AND MZF1B-C).
 RC TISSUE=Bone marrow; PubMed=10974541;
 RX MEDLINE=20432092; PubMed=10974541;
 RA Peterson M.J., Morris J.F.;
 RT "Human myeloid zinc finger gene MZF produces multiple transcripts and
 RT encodes a SCAN box protein";
 RL Gene 254:105-118 (2000).
 CC -!- FUNCTION: May be one regulator of transcriptional events during
 CC hemopoietic development.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=MZF1A; Synonyms=MZF1B;
 CC IsoId=P28698-1; Sequence=Displayed;
 CC Name=MZF1B-C;
 CC IsoId=P28698-2; Sequence=VSP_006889, VSP_006890;
 CC -!- TISSUE SPECIFICITY: Preferentially expressed in differentiating
 CC myeloid cells.
 CC -!- INDUCTION: By retinoic acid.
 CC -!- SIMILARITY: BELONGS TO THE KUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: Contains 1 SCAN box domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL; M38297; AAS5989.1; -;
 DR EMBL; AF055078; AAD55810.1; -;
 DR EMBL; AF055077; AAD55809.1; -;
 DR EMBL; AF161886; AAF80466.1; -;
 DR EMBL; AF161886; AAF80465.1; -;
 DR PIR; A40751; A40751.
 DR HSSP; P08047; 1SP2.
 DR TRANSFAC; T00529; -;
 DR GeneW; HGNC:13108; ZNF42.
 DR MIM; 194550; -;
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR InterPro; IPR003309; Treg SCAN.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF02023; SCAN; 1.
 DR Pfam; PF00096; zf_C2H2; 13.
 DR PRINTS; PR00048; ZINCFINGER.
 DR PRODOM; PD000003; Znf_C2H2; 6.
 DR SMART; SM00431; LER; 1.
 DR SMART; SM00355; ZNF_C2H2; 13.
 DR PROSITE; PS50804; SCAN_BOX; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.

DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 13.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat; Alternative splicing; Polymorphism.
 FT DOMAIN 44 125
 FT 310 321 ASP/GLU-RICH (ACIDIC).
 FT ZN_FING 356 378 C2H2-TYPE.
 FT ZN_FING 384 406 C2H2-TYPE.
 FT ZN_FING 412 434 C2H2-TYPE.
 FT ZN_FING 440 462 C2H2-TYPE.
 FT ZN_FING 463 484 GLY/PRO-RICH.
 FT DOMAIN 485 507
 FT ZN_FING 513 535 C2H2-TYPE.
 FT ZN_FING 541 563 C2H2-TYPE.
 FT ZN_FING 569 591 C2H2-TYPE.
 FT ZN_FING 597 619 C2H2-TYPE.
 FT ZN_FING 625 647 C2H2-TYPE.
 FT ZN_FING 653 675 C2H2-TYPE.
 FT ZN_FING 681 703 C2H2-TYPE.
 FT ZN_FING 709 731 C2H2-TYPE.
 FT VARSFING 1 249 Missing (in isoform MZF1B-C).
 FT VARSFING 250 257 /FTId=VSP_006889.
 FT VARSFING 257 /FTId=VSP_006890.
 FT VARSFING 331 331 I -> V (in dbSNP:4756).
 FT VARSFING 304 305 /FTId=VAR_014826.
 FT VARSFING 305 305 AL -> RV (in REF. 1).
 SQ SEQUENCE 734 AA; 82036 MW; 2BE7D69B18F29437 CRC64;
 Query Match 47.8%; Score 44; DB 1; Length 734;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 5 STRGRSGRSGXSG 18
 Db 337 STRGRSGRSGPSTG 350
 RESULT 11
 ID Y005 CAEEL STANDARD; PRT; 1232 AA.
 AC P34633;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Hypothetical protein ZK512.5 in chromosome III.
 GN ZK512.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berts M.,
 RA Bonfield J., Buron J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkes T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kersey J., Kirschen J., Laitster N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Snowken R.,
 RA Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38 (1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z23177; CA80146.1; -.
DR PIR: S40766; S40766.
DR WormPep; ZK512.5; CE00411.
KW Hypothetical protein.
SQ SEQUENCE 1232 AA; 134923 MW; 6DFC35D664AA8D6A CRC64;

Query Match 47.8%; Score 44; DB 1; Length 1232;
Best Local Similarity 64.3%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HOESTRGRSRGRSG 15
Db 110 HONSSRGPSGSG 123

RESULT 12
LSM4_CABEL STANDARD; PRT; 123 AA.
AC Q19958;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.
GN F32A5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Petodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds specifically to the 3'-terminal U-trace of U6
CC snRNA (by similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (by
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U20864; AAC4661.1; -.
DR PIR: T16234; T16234.
DR WormPep; F32A5.7; CE01277.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; LSM; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00551; Sm; 1.
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
KW RNA-binding.
SQ SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 123;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTRGRSRGRSGXSG 18
Db 93 REQSRGRGGRGGRG 108

```

```

RESULT 13
ID Y043 HALN1 STANDARD; PRT; 210 AA.
AC 09NH82;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Yng2543c.
GN YNG2543C.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.S., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Sprogon J.,
RA Swartzell S., Weir D., Hall U., Dahl T.A., Weir D.W.,
RA Lettner B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddock D.G., Jablonski P.E., Krebs W.P., Angvine C.M., Dale H.,
RA Isehnacker T.A., Peck R.F., Pohlechner M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RA "Genome sequence of Halobacterium species NRC-1";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- SIMILARITY: Contains 1 AMMECR1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005130; AAG20599.1; -.
DR PIR: C64404; C64404.
DR HAMAP; MF_00645; atypical; 1.
DR InterPro; IPR002733; DUF51.
DR Pfam; PF01871; AMMECR1; 1.
DR ProDom; PD009671; DUF51; 1.
DR TIGRFAMs; TIGR00296; TIGR00296; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 210 AA; 22964 MW; D4CBEOCEB4307EFF CRC64;

Query Match 46.7%; Score 43; DB 1; Length 210;
Best Local Similarity 61.1%; Pred. No. 76;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 4 ESTRGRSRGRSGXSGS 19
Db 55 ESTRGRRLRGCAHGS 72

RESULT 14
ID U2AG_MOUSE STANDARD; PRT; 239 AA.
AC U2AG_MOUSE;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE subunit) (U2 snRNP auxiliary factor small subunit).
GN U2AF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=small intestine;

```

RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Komno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
 RA Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schmitt L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita K., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RP [2]
 RP SEQUENCE OF 2-239 FROM N.A.
 RP TISSUE=Breast; tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish P.,
 RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Utschi T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalski D., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [1]
 RP FUNCTION: Plays a critical role in both constitutive and enhancer-
 dependent splicing by mediating protein-protein interactions and
 protein-RNA interactions required for accurate 3' splice site
 selection. Recruits U2 snRNP to the branch point. Directly
 mediates interactions between U2AF2 and proteins bound to the
 enhancers and thus may function as a bridge between U2AF2 and the
 enhancer complex to recruit it to the adjacent intron (By
 similarity).
 CC SUBUNIT: Heterodimer with U2AF2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- DOMAIN: The C-terminal SR-rich domain is required for
 interactions with SR proteins and the splicing regulators TRA and
 TR2, and the N-terminal domain is required for formation of
 the U2AF5/2AF5 heterodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the SR family of splicing factors.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AK008332; BAB25609.1; -
 DR EMBL; AK012849; BAB26511.1; -

DR EMBL; BC002184; AA02184.1; -
 DR MGD; MGI:988884; U2af1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR000571; Znf_CCH.
 DR Pfam; PF00076; itm; 1.
 DR Pfam; PF00642; zf-CCH; 2.
 DR SMART; SM00360; RRM; 1.
 DR SMART; SM00356; Znf_C3H1; 2.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 KW mRNA processing; mRNA splicing; Nuclear protein; Spliceosome;
 RN RNA-binding; Metal-binding; Zinc; Repeat; Zinc-finger.
 FT ZN_FING 13 41
 FT DOMAIN 65 147 RNA-BINDING (RRM).
 FT ZN_FING 149 173 C3H1-TYPE 2.
 FT DOMAIN 179 238 ARG/GLY/SER-RICH (RS DOMAIN).
 FT COMPLET 187 187 G->R (IN REF. 1; BAB25609).
 SQ SEQUENCE 239 AA; 27815 MW; DF944210581244d CAC64;
 Query Match 46.7%; Score 43; DB 1; Length 239;
 Best Local Similarity 44.4%; Pred. No. 8.7;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 1 SHQESTRGSRGSGXSG 18
 Db 199 SRERSRSDRGSGGCG 216
 RESULT 15
 U2AG HUMAN STANDARD; PRT; 240 AA.
 ID "U2AG HUMAN
 AC 001081;
 DT 01-APR-1993 (rel. 25, Created)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 15-MAR-2004 (rel. 43, Last annotation update)
 DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
 subunit) (U2 snRNP auxiliary factor small subunit).
 GN U2AF1 OR U2AF35.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 68-89 AND 126-151.
 RP TISSUE=Petal brain;
 RX MEDLINE=92409598; PubMed=1389271;
 RA Zhang W., Zamore P.D., Carmo-Fonseca M., Lamond A.I., Green M.R.,
 RT "Cloning and intracellular localization of the U2 small nuclear
 ribonucleoprotein auxiliary factor small subunit."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8769-8773(1992).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Tauden S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Paterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudch J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P.,
 RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Rameyer J., Beck A., Klages S., Hennig S., Rieseblum L., Dagand E.,
 RA Wehner S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehnach H., Reinhardt R., Yaspo M.-L.,
 RT "The DNA sequence of human chromosome 21."
 RL Nature 403:311-319(2000).
 RP [3]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Blatchenko L., Marzula K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Panse C.,
RA Raba S.S., Locuelli N.A., Peters G.J., Abramson R.D., Molligh S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy U., Helton E., Keltman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.W., Maria W.A.,
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP FUNCTION
RX MEDLINE=96249383; PubMed=8647433;
RA Zhu P., Maniatis T.;
RT "The splicing factor U2AF35 mediates critical protein-protein
RT interactions in constitutive and enhancer-dependent splicing.";
RL Genes Dev. 10:1356-1368(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 43-146 IN COMPLEX WITH U2AF2,
RP AND MUTAGENESIS OF TRP-134.
RX MEDLINE=21435806; PubMed=11551507;
RA Klempf C.L., Rodionova N.A., Green M.R., Buxley S.K.;
RT "A novel peptide recognition mode revealed by the X-ray structure of
RT a core U2AF35/U2AF65 heterodimer.";
RL Cell 106:595-605(2001).
CC -|- FUNCTION: Plays a critical role in both constitutive and enhancer-
CC dependent splicing by mediating protein-protein interactions and
CC protein-RNA interactions required for accurate 3' splice site
CC selection. Recruits U2 snRNP to the branch point. Directly
CC mediates interactions between U2AF2 and proteins bound to the
CC enhancers and thus may function as a bridge between U2AF2 and the
CC enhancer complex to recruit it to the adjacent intron.
CC -|- SUBUNIT: Heterodimer with U2AF2.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- DOMAIN: The C-terminal SR-rich domain is required for
CC interactions with SR proteins and the splicing regulators TRA and
CC TPA2, and the N-terminal domain is required for formation of
CC the U2AF1/U2AF2 heterodimer.
CC -|- SIMILARITY: Belongs to the SR family of splicing factors.
CC -|- SIMILARITY: Contains 2 C3H1-type zinc fingers.
CC -|- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M96982; AAA36619.1; -;
DR EMBL; AP001748; BAA95534.1; -;
DR EMBL; BC001177; AA01177.1; -;
DR EMBL; BC001923; AA01923.1; -;
DR PIR; A46179; A46179.
DR PDB; 1JMT; 1J-SEP-01.
DR Genew; HGNC:12453; U2AF1.
DR GK; O01081; -;
DR KIM; 191317; -;
DR GO; GO:0015030; C:capal body; TAS.
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.
DR GO; GO:0006397; F:mRNA processing; TAS.
DR GO; GO:0006371; P:mRNA splicing; TAS.
DR InterPro; IPR005054; RNA_rec_mot.
DR InterPro; IPR005571; Znf_CCH.

DR Pfam; PF00076; itm; 1.
DR Pfam; PF00642; zf-CCCH; 2.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00356; Znf_C3H1; 2.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
KW mRNA processing; mRNA_splicing; Nuclear protein; Spliceosome;
KW RNA-binding; Metal-binding; Zinc; Repeat; Zinc-finger; 3D-structure.
FT ZN FING 13 41
FT C3H1-TYPE 1. (RRM).
FT DOMAIN 65 147
FT ZN FING 149 173
FT DOMAIN 178 240
FT C3H1-TYPE 2.
FT ARG/GLY/SER-RICH (RS DOMAIN).
FT POLY-GLY.
FT MUTAGEN 134 134
FT W->A: DECREASES AFFINITY FOR UAF2 BY 3
FT ORDERS OF MAGNITUDE.
SQ SEQUENCE 240 AA; 27872 MW; 3DA130DC0B953F6 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 240;
Best Local Similarity 44.4%; Pred. No. 8.7;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTGRSRGRGSGXG 18
| : : | | | | |
DB 199 SRERRSRDRGRGGGGG 216

Search completed: September 28, 2004, 06:05:49
Job time : 8.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-5

Perfect score: 92

Sequence: 1 SHQESTRGSRGRSGXSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	797	4	Q16824
2	86	93.5	990	4	Q15206
3	86	93.5	1218	4	Q05331
4	84	91.3	798	4	Q9H4U3
5	84	91.3	1084	4	Q01212
6	75	81.5	591	4	Q01720
7	75	81.5	687	4	Q9H4U2
8	70	76.1	465	4	Q03838
9	69	75.0	322	4	Q75370
10	52	56.5	476	12	Q80890
11	48	52.2	188	10	Q8H8P6
12	48	52.2	336	10	Q943T8
13	48	52.2	820	12	Q69375
14	48	52.2	822	3	Q9P312
15	48	52.2	1229	3	Q862H1
16	47	51.1	52	16	Q98B32

17	47	51.1	373	10	Q23876	Q23876 oryza sativ
18	47	51.1	436	16	Q82C67	Q82C67 streptomyc
19	47	51.1	471	5	Q9UAY0	Q9UAY0 caenorhabd
20	47	51.1	571	13	Q8U1D0	Q8U1D0 oryzae lat
21	46.5	50.5	1150	5	Q9W0H4	Q9W0H4 arabidopsis
22	46	50.0	319	10	Q94J00	Q94J00 arabidopsis
23	46	50.0	330	5	Q7YXC2	Q7YXC2 cryptospori
24	46	50.0	467	10	Q61429	Q61429 arabidopsi
25	46	50.0	511	12	Q91332	Q91332 ceratophic
26	46	50.0	759	10	Q23135	Q23135 arabidopsi
27	46	50.0	810	10	Q94ES0	Q94ES0 pisum sativ
28	46	50.0	836	10	Q9SKZ9	Q9SKZ9 arabidopsi
29	46	50.0	1765	5	Q8SKX3	Q8SKX3 drosophila
30	46	50.0	2232	5	Q9VC96	Q9VC96 drosophila
31	45	48.9	398	5	Q9VX29	Q9VX29 drosophila
32	45	48.9	409	16	Q69574	Q69574 streptomyc
33	45	48.9	488	4	Q81UV8	Q81UV8 homo sapien
34	45	48.9	688	13	Q42375	Q42375 brachydanio
35	45	48.9	722	13	Q802W4	Q802W4 brachydanio
36	45	48.9	899	10	Q9FF14	Q9FF14 arabidopsi
37	45	48.9	1142	13	Q752V2	Q752V2 xenopus lae
38	45	48.9	1483	5	Q9BJW3	Q9BJW3 plasmodium
39	44.5	48.4	573	5	Q81GR2	Q81GR2 drosophila
40	44.5	48.4	682	5	Q44553	Q44553 caenorhabdi
41	44.5	48.4	697	13	Q72XJ0	Q72XJ0 xenopus lae
42	44.5	48.4	1496	10	Q22175	Q22175 arabidopsi
43	44.5	48.4	1910	5	Q9V768	Q9V768 drosophila
44	44	47.8	222	4	Q96FA0	Q96FA0 homo sapien
45	44	47.8	229	11	Q9JW93	Q9JW93 mus musculu

ALIGNMENTS

RESULT 1

ID Q16824 PRELIMINARY; PRT; 797 AA.

AC Q16824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Seibert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).

DR EMBL; M60502; AAA63248.1; -
DR GO; GO:0005196; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF03516; Flaggrin; 4.
DR PRINTS; PR00487; FLAGGRIN.
FT NON TER 1

SQ SEQUENCE 797 AA; 85176 MW; 6056184763BDA868 CRC64;

Query Match 97.8%; Score 90; DB 4; Length 797;
Best Local Similarity 94.7%; Pred. No. 4.3e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTRGSRGRSGXSGS 19
|||
Db 427 SHQESTRGSRGRSGXSGS 445

RESULT 2

```

015206
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
EX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814]";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
EX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AA63244.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10FA91991 CRC64;

Query Match 93.5%; Score 86; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 2.4e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 551 SHOESTRGRSGRSGSGS 569

RESULT 3
005331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
EX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipkev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
DE -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DE DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
DE TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M69943; AAA36487.1; -.
DR PIR; A48118; A48118.
DR HSSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005505; F:calcium ion binding; TAS.
DR GO; GO:0030154; F:cell differentiation; NAS.
DR GO; GO:0008151; F:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CAPS S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KM Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 93.5%; Score 86; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 773 SHOESTRGRSGRSGSGS 791

RESULT 4
Q9H4U3 PRELIMINARY; PRT; 798 AA.
AC Q9H4U3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 91.3%; Score 84; DB 4; Length 798;
Best Local Similarity 89.5%; Pred. No. 4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSGRSGSGS 19
DB 428 SHOESTRGRSGRSGSGS 446

RESULT 5
Q01212 PRELIMINARY; PRT; 1084 AA.
ID Q01212

```

AC 001212; 003840;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9106347; PubMed=2248957;
 RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814].";
 RT 11;30(23):5814.";
 RL Biochemistry 29:9432-9440(1990).
 DR EMBL; M60503; AAA63243.1; JOINED.
 DR EMBL; M60501; AAA63243.1; JOINED.
 DR GO; GO:0005882; C:intermediate filament; NAS.
 DR GO; GO:0005198; F:structural molecule activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PRO0487; Filaggrin.
 FT NON TER
 SQ SEQUENCE 1084 AA; 115271 MW; 80C46408BD5A362D CRC64;

Query Match 91.3%; Score 84; DB 4; Length 1084;
 Best Local Similarity 89.5%; Pred. No. 5.5e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQSTRGRSGRSGSGS 19
 Db 64 SHQSTRGRSGRSGSGS 82

RESULT 6
 Q01720
 ID 001720 PRELIMINARY; PRT; 591 AA.
 AC 001720;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Filaggrin precursor (PROFILAGGRIN) (Fragment).
 GN FIG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Presland R.B., Haycock P.V., Fleckman P., Nirmusksiri W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 organization and identification of an S-100-like calcium binding
 domain at the amino terminus.";
 RT J. Biol. Chem. 267:23772-23781(1992).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 324 AA. THE PRECURSOR IS DEPOSITED AS KERSOTHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTOLOGICALLY CLEAVED.
 CC -1- POLYMORPHISMS: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; L01089; AAA60177.1; -;
 DR EMBL; L01090; AAA60176.1; -;

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; Filaggrin.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KW Polymorphism.
 FT PROPEP 1 293
 FT CHAIN 294 467
 FT PROPEP 468 474
 FT CHAIN 475 >591
 FT CA BIND 19 32
 FT CA BIND 62 73
 FT NON TER 591 591
 SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 591;
 Best Local Similarity 84.2%; Pred. No. 0.00084;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHQSTRGRSGRSGSGS 19
 Db 449 SHQSTRGRSGRSGSGS 467

RESULT 7
 Q09402
 ID 09402 PRELIMINARY; PRT; 687 AA.
 AC 09402;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Df14N1.1.1 (Profilaggrin 5' end) (Fragment).
 GN FIG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL356504; CAC13172.1; -;
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR003303; Filaggrin.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; Filaggrin.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 FT NON TER 687 687
 SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 81.5%; Score 75; DB 4; Length 687;
 Best Local Similarity 84.2%; Pred. No. 0.00098;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGSGSGS 19
DB 449 SHOESTRGSRGSGSGS 467

RESULT 8

ID Q03838 PRELIMINARY; PRT; 465 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE FILAGRIN (PROFILAGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

FT NON_TER 1 322

RT "Organizational, structure, and polymorphisms of the human profilaggrin gene."

RL Biochemistry 29:9432-9440(1990).

RN [2]

RP REVISIONS.

RA MEDLINE=91255199; PubMed=2043621;

RT "Organizational, structure, and polymorphisms of the human profilaggrin gene."

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE

CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN

CC REPEATS.

CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM

CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin

CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

CC EMBL; M60499; AAA63246.1; -

CC GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin.3.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1 465

SQ SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

Query Match 76.1%; Score 70; DB 4; Length 465;

Best Local Similarity 78.9%; Pred. No. 0.0042; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 4;

QY 1 SHOESTRGSRGSGSGS 19

DB 227 SHOESTRGSRGSGSGS 245

RESULT 9

ID Q75370 PRELIMINARY; PRT; 322 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Epidermal filaggrin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=99101527; PubMed=9886436;

RA Girbal-Neuhausser E., Durieux J.J., Arnaud M., Dalbon P., Sebbaq M.,

RA Vincent C., Simon M., Sersu T., Masson-Bessiere C.,

RA Jolivet-Reynaud C., Jolivet M., Serre G.,

RT "The epitopes targeted by the rheumatoid arthritis-associated

RT anti-filaggrin autoantibodies are posttranslationally generated on

RT various sites of (pro)filaggrin by deamination of arginine residues."

RL J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC23559.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin.2.

DR PRINTS; PR00487; FILAGRIN.

FT NON_TER 1 322

SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

Query Match 75.0%; Score 69; DB 4; Length 322;

Best Local Similarity 77.8%; Pred. No. 0.0041;

Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGSGSGS 18

DB 305 SHOESTRGSRGSGSGS 322

RESULT 10

ID Q80890 PRELIMINARY; PRT; 476 AA.

AC Q80890;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE EBNA1.

OS Herpesvirus papio.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gammaherpesvirinae; Lymphocryptovirus.

OX NCBI_TaxID=10394;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=594-S;

RA MEDLINE=96400106; PubMed=8806482;

RA Yates J.L., Camilo S.M., Ali S., Ying A.,

RT "Comparison of the EBNA1 proteins of Epstein-Barr virus and

RT herpesvirus papio in sequence and function."

RL Virology 222:1-13(1996).

DR EMBL; U23857; AAA66373.1; -

DR HSP; P03211; IVH1.

DR GO; GO:0042025; C:host cell nucleus; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR InterPro: IPR004186; EBNA1.

DR InterPro: IPR009021; Viral_DNA_Bd.

DR Pfam; PF02905; EBNA1.1;

SQ SEQUENCE 476 AA; 46082 MW; 3F598F27919F7DA9 CRC64;

Query Match 56.5%; Score 52; DB 12; Length 476;

Best Local Similarity 71.4%; Pred. No. 3.5;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 11

ID Q8H8P6 PRELIMINARY; PRT; 188 AA.

AC Q8H8P6;

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

```

DE 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Hypothetical protein.
GN OSUNBA0057E11.19.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Brell C.R., Yuan Q., Guyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Heiso J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uterback T.T., Quackendush J.,
RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.C., Quackendush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0057E11 genomic sequence.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC091234; AAN64459.1; -
DR InterPro: IPR006552; DUF834.
DR Pfam: PF05754; DUF834; 1.
KM Hypothetical protein.
SQ SEQUENCE 188 AA; 19599 MW; DCSA0096F51192B5 CRC64;

Query Match 52.2%; Score 48; DB 10; Length 188;
Best Local Similarity 60.0%; Pred. No. 5;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 ESTGRSRGRSGXSG 18
Db 168 EQIRGRGRGSGXSG 182

RESULT 12
Q94378 PRELIMINARY; PRT; 336 AA.
ID Q94378;
AC Q94378;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE B1158F07.1 protein (OJ1008_F01.12 protein).
GN B1158F07.1 OR OJ1008_F01.12.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530; 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=O. sativa; STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
RT clone:OJ1008_F01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=O. sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, BAC
RT clone:OJ1008_F01.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003208; BAB64590.1; -
DR EMBL: AF004320; BAC00705.1; -
DR Gramene; Q94378; -
DR InterPro: IPR008552; DUF834.
DR Pfam: PF05754; DUF834; 1.
SQ SEQUENCE 336 AA; 35999 MW; 5080CDAFAC7B9B88 CRC64;

Query Match 52.2%; Score 48; DB 10; Length 336;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 ESTGRSRGRSGXSG 18
Db 316 EQIRGRGRGSGXSG 330

RESULT 13
Q69375 PRELIMINARY; PRT; 820 AA.
ID Q69375;
AC Q69375;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, last annotation update)
DE Putative UR25 protein.
OS Mouse cytomegalovirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10366;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KL81;
RX MEDLINE=94233727; PubMed=7513920;
RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;
RT "Identification and characterization of a murine cytomegalovirus gene
RT with homology to the UR25 open reading frame of human
RT cytomegalovirus.";
RL Virology 200;643-650(1994).
DR EMBL: U02500; AAL19449.1; -
DR InterPro: IPR006731; Herpes_PP85.
DR Pfam: PF04637; Herpes_PP85; 1.
SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;

Query Match 52.2%; Score 48; DB 12; Length 820;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QESTGRSRGRS 14
Db 186 QSTGRSRGRS 197

RESULT 14
Q9P312 PRELIMINARY; PRT; 822 AA.
ID Q9P312;
AC Q9P312;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12P1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohnsels J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL390091; CAB98213.1; -
DR PIR; T51049; T51049.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR InterPro: IPR005054; RNA_rec_mot.
DR InterPro: IPR005120; Smg-4_UPF3.
DR Pfam: PF00076; trm; 1.
DR Pfam: PF03467; Smg4_UPF3; 1.
DR SMART; SM00360; RRM_1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.

```

SO SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;
 Query Match 52.2%; Score 48; DB 3; Length 822;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTRGRSGRSGSGS 19
 :|||:|||||:
 Db 414 RESASGRTRGRGRGT 430

RESULT 15

Q86ZHL PRELIMINARY; PRT; 1229 AA.
 ID Q86ZHL
 AC Q86ZHL; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Related to cell division cycle 2-related protein kinase 7.
 GN 7F4.060.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RX [1]
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nykatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BX294020; CAD70910.1; -;
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: GO:0000910; P:cytokinesis; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR InterPro: IPR008271; Ser Thr kinase.
 DR InterPro: IPR01245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD00001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Cell division; Kinase.
 SQ SEQUENCE 1229 AA; 136776 MW; 43BE612656FF73E4 CRC64;

Query Match 52.2%; Score 48; DB 3; Length 1229;
 Best Local Similarity 57.9%; Pred. No. 42;
 Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 SHOESTGRSGRSGSGS 19
 :|||:|||||:
 Db 242 SHHRRRRSDKRRSGRSRS 260

Search completed: September 28, 2004, 06:12:44
 Job time : 42.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 / Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-5
Perfect score: 92
Sequence: 1 SHQSTRGRSRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	2	AAW61509 Peptide c
2	90	97.8	19	2	AAW61514 Peptide c
3	86	93.5	19	2	AAW61517 Peptide c
4	86	93.5	1457	5	ABB97605 Novel hum
5	85	92.4	19	2	AAW61516 Peptide c
6	84	91.3	19	2	AAW61506 Peptide c
7	84	91.3	19	2	AAW61508 Peptide c
8	84	91.3	19	2	AAW61515 Peptide c
9	84	91.3	19	2	AAW61507 Peptide c
10	84	91.3	19	2	AAW61513 Peptide c
11	84	91.3	19	2	AAW61505 Peptide c
12	78	84.8	19	2	AAW61512 Peptide c
13	78	84.8	19	2	AAW61511 Peptide c
14	78	84.8	19	2	AAW61510 Peptide c
15	75	81.5	21	2	AAW61520 Peptide X
16	75	81.5	22	4	AAE07235 IGP1546 p
17	64	69.6	330	2	AAW22956 Human fil
18	64	69.6	330	2	AAW22955 Human fil
19	64	69.6	330	2	AAW22957 Human fil
20	60	65.2	330	2	AAW22954 Human fil
21	56	60.9	477	6	ABO07142 Novel hum
22	51	55.4	1711	4	AAW79819 Human pro
23	51	55.4	1951	4	AAW78835 Human pro
24	49	53.3	641	4	ABG19110 Novel hum
25	48	52.2	772	4	AAW64573 Human Mre

26	47	51.1	184	3	AAW57041 Human pro
27	46.5	50.5	1150	4	ABW59129 Arabidops
28	46	50.0	306	3	AAW06829 Arabidops
29	46	50.0	467	5	ABW92835 Arabidops
30	46	50.0	532	4	ABW67173 Arabidops
31	45	48.9	99	4	AAW62331 Propionib
32	45	48.9	99	6	AAW58850 Propionib
33	45	48.9	398	4	ABW58306 Drosophila
34	45	48.9	883	6	AAW43571 Propionib
35	45	48.9	883	6	AAW40090 Propionib
36	45	48.9	1703	6	ABW52985 Protein s
37	44.5	48.4	1945	4	ABW64947 Drosophila
38	44	47.8	131	4	AAW06599 Human pol
39	44	47.8	141	7	AAW09023 Novel pro
40	44	47.8	157	3	AAW41022 Zeta may
41	44	47.8	207	7	AAW72443 Human end
42	44	47.8	222	4	AAW93652 Human pol
43	44	47.8	229	7	AAW59825 Rat prote
44	44	47.8	229	7	AAW59833 Rat prote
45	44	47.8	229	7	AAW59837 Rat Prote

ALIGNMENTS

RESULT 1	AAW61509	AAW61509 standard; peptide; 19 AA.
ID	AAW61509	standard; peptide; 19 AA.
XX	XX	
AC	AAW61509;	
XX	XX	
DT	26-OCT-1998	(first entry)
XX	XX	
DE	Peptide cfc5, based on cDNA of a profilaggrin repeat.	
XX	XX	
KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
KM	solid phase synthesis; peptide amide; polyclonal antibody;	
KX	monoclonal antibody.	
XX	XX	
OS	Synthetic.	
OS	Homo sapiens.	
XX	XX	
FH	Key	Location/Qualifiers
FT	Modified-site	16 /note="Citruiline"
XX	XX	
PN	WO9822503-A2.	
XX	XX	
PD	28-MAY-1998.	
XX	XX	
PF	14-NOV-1997;	97WO-NL000624.
XX	XX	
PR	15-NOV-1996;	96NL-01004539.
XX	XX	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
XX	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
PI	Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;	
XX	WPI, 1998-398613/34.	
XX	XX	
DR	Peptide derived from an antigen recognised by autoantibodies - is	
XX	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
XX	used in diagnosis of the disease.	
PT	Disclosure; Page 6; 19pp; English.	
PS	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of	
XX	the profilaggrin antigen which is recognised by autoantibodies from	
CC	patients with rheumatoid arthritis (RA). This peptide is reactive with a	
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	The peptides were created by using standard solid phase synthesis, which	
CC	produced them as peptide amides. These sequences may be used in the	

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 2

AAW61514
 ID AAW61514 standard; peptide: 19 AA.

AAW61514;

26-OCT-1998 (first entry)

Peptide cf, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 solid phase synthesis; peptide amide; polyclonal antibody;
 monoclonal antibody.

Synthetic.

Homo sapiens.

WO9822503-A2.

28-MAY-1998.

14-NOV-1997; 97WO-NL000624.

15-NOV-1996; 96NL-01004539.

(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

(TEWE-) STICHTING TECH WETENSCHAPPEN.

Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;

WPI; 1998-398613/34.

Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA;

Query Match 97.8%; Score 90; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 1.7e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 3

AAW61517
 ID AAW61517 standard; peptide: 19 AA.

AAW61517;

26-OCT-1998 (first entry)

Peptide cfQ, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 solid phase synthesis; peptide amide; polyclonal antibody;
 monoclonal antibody.

Synthetic.

Homo sapiens.

WO9822503-A2.

28-MAY-1998.

14-NOV-1997; 97WO-NL000624.

15-NOV-1996; 96NL-01004539.

(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

(TEWE-) STICHTING TECH WETENSCHAPPEN.

Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;

WPI; 1998-398613/34.

Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA;

Query Match 93.5%; Score 86; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.5e-07;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTGRSRGRSGSGS 19

RESULT 4

ABB97605
 ID ABB97605 standard; protein: 1467 AA.

ABB97605;

27-JUN-2002 (first entry)

Novel human protein SEQ ID NO: 873.

Human; antihaemic; vulnery; antiinflammatory; immunomodulator;
 antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 neuroprotective; antiparkinsonian; protein therapy; BSR;
 expressed sequence tag.

Homo sapiens.

```

XX PN W0200222660-A2.
XX PD 21-MAR-2002.
XX PF 10-SEP-2001; 2001WO-US026015.
XX PR 11-SEP-2000; 2000US-00659671.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Wehrman T, Dermanac RT;
XX DR WPI; 2002-282408/33.
XX DR N-PSDB; ABN32791.
XX PT An isolated polynucleotide for treating diseases associated with its
XX PT encoded polypeptide such as cancer and multiple sclerosis.
XX PS Example 2; SEQ ID NO 873; 509pp; English.
XX CC The present invention provides the protein and coding sequences of 444
XX CC novel human proteins. These were isolated from expressed sequences tags
XX CC (ESTs). They can be used to stimulate cell growth, to regulate
XX CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
XX CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
XX CC rheumatoid arthritis, and to treat nervous system disorders e.g.
XX CC Parkinson's disease. The present sequence is a protein of the invention
XX SQ Sequence 1467 AA;

Query Match 93.5%; Score 86; DB 5; Length 1467;
Best Local Similarity 89.5%; Pred. No. 6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGSRGSGS 19
   ||||| ||||| ||||| |||||
Db 773 SHOESTRGSRGSRGSGS 791

RESULT 5
AAW61516
ID AAW61516 standard; peptide; 19 AA.
XX AAW61516;
XX AC
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cFE, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclonal antibody.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN W09822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

```

```

XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX SQ Sequence 19 AA;

Query Match 92.4%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGSRGSGS 19
   ||||| ||||| ||||| |||||
Db 1 SHOESTRGSRGSRGSGS 19

RESULT 6
AAW61506
ID AAW61506 standard; peptide; 19 AA.
XX AAW61506;
XX AC
XX DT 26-OCT-1998 (first entry)
XX DE Peptide cFE2, based on cDNA of a profilaggrin repeat.
XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclonal antibody.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 9 /note= "Citruilline"
XX PN W09822503-A2.
XX PD 28-MAY-1998.
XX PF 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the profilaggrin antigen which is recognised by autoantibodies from

```

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTRGRSRGRSGSGS 19
DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 7
AAW61508
ID AAW61508 standard; peptide; 19 AA.
XX
AC AAW61508;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfc4, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 13
FT /note= "Citrulline"
XX
FN WO9822503-A2.
XX
PD 28-MAY-1998.
XX
PF 14-NOV-1997; 97WO-NL000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX
DR WPI; 1998-398613/34.
XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
SQ Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTRGRSRGRSGSGS 19
DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 8
AAW61515
ID AAW61515 standard; peptide; 19 AA.
XX
AC AAW61515;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfa, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
FN WO9822503-A2.
XX
PD 28-MAY-1998.
XX
PF 14-NOV-1997; 97WO-NL000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX
DR WPI; 1998-398613/34.
XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19pp; English.
XX
SQ Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;

Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTRGRSRGRSGSGS 19
DB 1 SHOESTRGRSRGRSGSGS 19
RESULT 9
AAW61507
ID AAW61507 standard; peptide; 19 AA.
XX
AC AAW61507;
XX
DT 26-OCT-1998 (first entry)
XX

```

DE Peptide cfc3, based on cDNA of a proflilagrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KM solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclinal antibody.
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Modified-site 11
FT /note= "Citruiline"
XX
XX W09822503-A2.
PN
XX
XX 28-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-NL000624.
PF
XX
XX 15-NOV-1996; 96NL-01004539.
PR
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
DR
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX PS Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the proflilagrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclinal antibodies
XX
XX SQ Sequence 19 AA;
XX
XX Query Match 91.3%; Score 84; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTRGSRGSGSGS 19
XX 1 SHOESTRGSRGSGSGS 19
XX
XX Db
XX
XX RESULT 10
XX AAW61513
XX ID AAW61513 standard; peptide; 19 AA.
XX
XX AC AAW61513;
XX
XX 26-OCT-1998 (first entry)
XX
XX DE Peptide cfc9, based on cDNA of a proflilagrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclinal antibody.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7
XX

```

```

FT /note= "Citruiline"
FT Modified-site 16
FT /note= "Citruiline"
XX
XX W09822503-A2.
PN
XX
XX 28-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-NL000624.
PF
XX
XX 15-NOV-1996; 96NL-01004539.
PR
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
DR
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX PS Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the proflilagrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclinal antibodies
XX
XX SQ Sequence 19 AA;
XX
XX Query Match 91.3%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 1.6e-06;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTRGSRGSGSGS 19
XX 1 SHOESTRGSRGSGSGS 19
XX
XX Db
XX
XX RESULT 11
XX AAW61505
XX ID AAW61505 standard; peptide; 19 AA.
XX
XX AC AAW61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX DE Peptide cfc1, based on cDNA of a proflilagrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclinal antibody.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7
XX /note= "Citruiline"
XX
XX W09822503-A2.
PN
XX
XX 28-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-NL000624.
PF
XX
XX 15-NOV-1996; 96NL-01004539.
PR

```

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij JMW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA:
Query Match 91.3%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHQESTGRGRGRSGSGS 19
DB 1 SHQESTXGRGRGRSGSGS 19
RESULT 12
AAM61512
ID AAM61512 standard; peptide; 19 AA.
XX
XX AAM61512;
AC
XX
XX 26-OCT-1998 (first entry)
DT
XX
XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
DE
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citnulline"
FT Modified-site 13
FT Modified-site 13 /note= "Citnulline"
FT
XX
XX W09822503-A2.
PN
XX
XX 28-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-NL000624.
PF
XX
XX 15-NOV-1996; 96NL-01004539.
PR
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij JMW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA:
Query Match 84.8%; Score 76; DB 2; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SHQESTGRGRGRSGSGS 19
DB 1 SHQESTXGRGRGRSGSGS 19
RESULT 13
AAM61511
ID AAM61511 standard; peptide; 19 AA.
XX
XX AAM61511;
AC
XX
XX 26-OCT-1998 (first entry)
DT
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
DE
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citnulline"
FT Modified-site 11
FT Modified-site 11 /note= "Citnulline"
FT
XX
XX W09822503-A2.
PN
XX
XX 28-MAY-1998.
PD
XX
XX 14-NOV-1997; 97WO-NL000624.
PF
XX
XX 15-NOV-1996; 96NL-01004539.
PR
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij JMW, Schellekens GA, Raats JMH, Hoet RMA;
PI WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC

```

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.4e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGRSGSGS 19
 |||||
 1 SHOESTXGSRGRSGSGS 19

DB

RESULT 14

AAW61510
 ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

AC 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citruilline"

FT Modified-site 9 /note= "Citruilline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 84.8%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.4e-05;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTRGSRGRSGSGS 19
 |||||
 1 SHOESTXGSRGRSGSGS 19

DB 1 SHOESTXGSRGRSGSGS 19

RESULT 15

AAW61520
 ID AAW61520 standard; peptide; 21 AA.

XX AAW61520;

AC 26-OCT-1998 (first entry)

XX Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note= "Citruilline"

XX W09822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 81.5%; Score 75; DB 2; Length 21;
 Best Local Similarity 83.3%; Pred. No. 4.7e-05;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HOESTRGSRGRSGSGS 19
 |||||
 4 HOESTXGSRGRSGSGS 21

DB 4 HOESTXGSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:15
 Job time : 56.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2004, 06:15:16 Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-5

Perfect score: 92

Sequence: 1 SHQESTRGRSGRSGXSGS 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	97.8	19	9	US-09-308-150-5
2	90	97.8	19	9	US-09-308-150-11
3	86	93.5	19	9	US-09-308-150-14
4	85	92.4	19	9	US-09-308-150-13
5	84	91.3	19	9	US-09-308-150-1
6	84	91.3	19	9	US-09-308-150-2
7	84	91.3	19	9	US-09-308-150-3
8	84	91.3	19	9	US-09-308-150-4
9	84	91.3	19	9	US-09-308-150-9
10	84	91.3	19	9	US-09-308-150-12
11	78	84.8	19	9	US-09-308-150-6
12	78	84.8	19	9	US-09-308-150-7
13	78	84.8	19	9	US-09-308-150-8
14	75	81.5	21	9	US-09-308-150-10
15	75	81.5	22	9	US-09-747-029A-22

16	56	60.9	477	15	US-10-161-927-62	Sequence 62, Appl
17	52	56.5	78	16	US-10-437-963-203073	Sequence 203073,
18	49	53.3	130	16	US-10-437-963-174101	Sequence 174101,
19	49	53.3	1071	16	US-10-437-963-118389	Sequence 118389,
20	48	52.2	1223	16	US-10-767-701-32436	Sequence 32436, A
21	48	52.2	172	12	US-10-424-599-215003	Sequence 215003,
22	48	52.2	172	12	US-10-425-114-39730	Sequence 39730, A
23	48	52.2	188	12	US-10-424-599-157696	Sequence 157696,
24	48	52.2	211	14	US-10-002-631C-268	Sequence 268, App
25	47	51.1	772	15	US-10-369-493-6474	Sequence 6474, Ap
26	47	51.1	184	9	US-09-925-300-1619	Sequence 1619, Ap
27	47	51.1	372	16	US-10-437-963-143403	Sequence 143403,
28	47	51.1	436	14	US-10-156-761-13022	Sequence 13022, A
29	47	51.1	492	16	US-10-437-963-120908	Sequence 120908,
30	47	51.1	511	16	US-10-437-963-188040	Sequence 188040,
31	47	51.1	563	16	US-10-437-963-198539	Sequence 198539,
32	47	51.1	838	16	US-10-437-963-146503	Sequence 146503,
33	47	51.1	854	12	US-10-425-114-57838	Sequence 57838, A
34	46	50.0	16	10	US-09-791-524-150	Sequence 150, App
35	46	50.0	142	16	US-10-767-701-49555	Sequence 49555, A
36	46	50.0	178	16	US-10-437-963-142959	Sequence 142959,
37	46	50.0	231	16	US-10-437-963-179346	Sequence 179346,
38	46	50.0	345	12	US-10-425-114-71489	Sequence 71489, A
39	46	50.0	461	16	US-10-437-963-185969	Sequence 185969,
40	46	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
41	46	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
42	45	48.9	49	12	US-10-424-599-187406	Sequence 187406,
43	45	48.9	150	16	US-10-767-701-52043	Sequence 52043, A
44	45	48.9	181	16	US-10-437-963-164001	Sequence 164001,
45	45	48.9	580	12	US-10-425-114-70952	Sequence 70952, A

ALIGNMENTS

RESULT 1

US-09-308-150-5

Sequence 5, Application US/09308150

Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Walthers Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoelt, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 19

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline

US-09-308-150-5

Query Match 97.8%; Score 90; DB 9; Length 19;

Best Local Similarity 100.0%; Pred. No. 4.2e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 2

US-09-308-150-11
 ; Sequence 11, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-11

Query Match 97.8%; Score 90; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 4.2e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 3

US-09-308-150-14
 ; Sequence 14, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 19

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:Derived from
 OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-14

Query Match 93.5%; Score 86; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.6e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 4

US-09-308-150-13
 ; Sequence 13, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
 ; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
 ; CURRENT FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: PCT/NL97/00624
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: NL1004539
 ; PRIOR FILING DATE: 1996-11-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:Derived from
 ; OTHER INFORMATION: known cDNA sequences of human profilaggrin
 US-09-308-150-13

Query Match 92.4%; Score 85; DB 9; Length 19;
 Best Local Similarity 89.5%; Pred. No. 2.2e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTRGRSRGRSGSGS 19
 DB 1 SHOESTRGRSRGRSGSGS 19

RESULT 5

US-09-308-150-1
 ; Sequence 1, Application US/09308150
 ; Patent No. US20020137092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
 ; APPLICANT: Schellekens, Gerardus Antonius
 ; APPLICANT: Raats, Jozef Maria Hendrik
 ; APPLICANT: Hoeft, Rene Michael Antonius
 ; APPLICANT: Stichting Scheikundig Onderzoek Nederland
 ; APPLICANT: Stichting voor de Technische Wetenschappen
 ; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
 ; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
 ; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```

; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGRSGSGS 19
Db 1 SHOESTRGSRGRSGSGS 19

```

```

RESULT 6
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Scheellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGRSGSGS 19
Db 1 SHOESTRGSRGRSGSGS 19

```

```

RESULT 7
US-09-308-150-3

```

```

; Sequence 3, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Scheellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

```

```

Query Match          91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 SHOESTRGSRGRSGSGS 19
Db 1 SHOESTRGSRGRSGSGS 19

```

```

RESULT 8
US-09-308-150-4
; Sequence 4, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Scheellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline

```

US-09-308-150-4

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 1 SHOESTGRSGRSGSGS 19

RESULT 9

US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoel, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 91.3%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 3e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 1 SHOESTGRSGRSGSGS 19

RESULT 10

US-09-308-150-12
; Sequence 12, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoel, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624

; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-12

Qy 1 SHOESTGRSGRSGSGS 19
Db 1 SHOESTGRSGRSGSGS 19

RESULT 11

US-09-308-150-6
; Sequence 6, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoel, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00021;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTGRSGRSGSGS 19
Db 1 SHOESTGRSGRSGSGS 19

RESULT 12

US-09-308-150-7
; Sequence 7, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius

```

/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoeft, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ APPLICANT: Stichting voor de Technische Wetenschappen
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00021;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 1 SHOESTRGRSRGRSGXSGS 19
DB      1 SHOESTRGRSRGRSGXSGS 19

```

```

RESULT 13
US-09-308-150-8
/ Sequence 8, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Schellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoeft, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 19
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      84.8%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 0.00021;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 1 SHOESTRGRSRGRSGXSGS 19
DB      1 SHOESTRGRSRGRSGXSGS 19

```

```

RESULT 14
US-09-308-150-10
/ Sequence 10, Application US/09308150
/ Patent No. US20020137092A1
/ GENERAL INFORMATION:
/ APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
/ APPLICANT: Schellekens, Gerardus Antonius
/ APPLICANT: Raats, Jozef Maria Hendrik
/ APPLICANT: Hoeft, Rene Michael Antonius
/ APPLICANT: Stichting Scheikundig Onderzoek Nederland
/ TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
/ TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
/ TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
/ FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
/ CURRENT APPLICATION NUMBER: US/09/308,150
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: PCT/NL97/00624
/ PRIOR FILING DATE: 1997-11-14
/ PRIOR APPLICATION NUMBER: NL 1004539
/ PRIOR FILING DATE: 1996-11-15
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 21
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Derived from
/ OTHER INFORMATION: known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
/ NAME/KEY: DISULFID
/ LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      81.5%; Score 75; DB 9; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00063;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 2 HOESTRGRSRGRSGXSGS 19
DB      4 HOESTRGRSRGRSGXSGS 21

```

```

RESULT 15
US-09-747-029A-22
/ Sequence 22, Application US/09747029A
/ Patent No. US20020143143A1
/ GENERAL INFORMATION:
/ APPLICANT: Union, Am
/ APPLICANT: Moereels, Henri
/ APPLICANT: Wehens, Lydie
/ TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
/ TITLE OF INVENTION: RHEUMATOID ARTHRITIS
/ FILE REFERENCE: 11362.003INPUS00 INNS:031
/ CURRENT APPLICATION NUMBER: US/09/747,029A
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: EP 00870195.5
/ PRIOR FILING DATE: 2000-09-08
/ PRIOR APPLICATION NUMBER: EP 99870280.7
/ PRIOR FILING DATE: 1999-12-21
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 22
/ LENGTH: 22
/ TYPE: PRT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-5

Perfect score: 92

Sequence: 1 SHQESTGRSGRSGXSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	51.1	629	4	US-09-252-991A-22901
2	44	47.8	157	4	US-09-252-991A-27836
3	44	47.8	653	4	US-09-513-057C-13
4	43	46.7	159	4	US-09-252-991A-27107
5	43	46.7	581	4	US-09-252-991A-25519
6	43	46.7	722	4	US-09-984-890-4
7	43	46.7	724	4	US-09-984-890-2
8	43	46.7	745	4	US-09-523-849-36
9	43	46.7	2237	1	US-08-354-973-1
10	42.5	46.2	398	4	US-09-252-991A-19488
11	42	45.7	105	4	US-09-252-991A-32394
12	42	45.7	349	4	US-09-252-991A-27336
13	42	45.7	523	4	US-09-252-991A-21124
14	41	44.6	70	2	US-08-724-548-54
15	41	44.6	70	3	US-07-978-674B-54
16	41	44.6	154	4	US-09-252-991A-32846
17	41	44.6	169	3	US-09-342-084-6
18	41	44.6	204	4	US-09-252-991A-21317
19	41	44.6	205	4	US-09-252-991A-18529
20	41	44.6	255	4	US-09-252-991A-19598
21	41	44.6	357	4	US-09-252-991A-17216
22	41	44.6	409	4	US-09-489-039A-14217
23	41	44.6	448	4	US-09-252-991A-24066
24	41	44.6	482	4	US-09-252-991A-16654
25	41	44.6	719	4	US-09-252-991A-28576
26	41	44.6	722	4	US-08-817-832B-32
27	40.5	44.0	257	4	US-09-252-991A-22204

28	40.5	44.0	703	3	US-08-910-925-4	Sequence 4, Appli
29	40	43.5	134	4	US-09-252-991A-18314	Sequence 18314, A
30	40	43.5	170	4	US-09-252-991A-26306	Sequence 26306, A
31	40	43.5	237	4	US-09-252-991A-23304	Sequence 23304, A
32	40	43.5	239	4	US-09-252-991A-28985	Sequence 28985, A
33	40	43.5	255	3	US-09-342-084-11	Sequence 11, Appl
34	40	43.5	274	4	US-09-252-991A-16696	Sequence 16696, A
35	40	43.5	286	4	US-09-252-991A-22717	Sequence 22717, A
36	40	43.5	294	4	US-09-252-991A-21947	Sequence 21947, A
37	40	43.5	326	4	US-09-252-991A-20075	Sequence 20075, A
38	40	43.5	363	4	US-09-252-991A-25132	Sequence 25132, A
39	40	43.5	406	4	US-09-252-991A-19857	Sequence 19857, A
40	40	43.5	421	4	US-09-252-991A-32316	Sequence 32316, A
41	40	43.5	478	4	US-09-252-991A-25411	Sequence 25411, A
42	40	43.5	480	4	US-09-252-991A-22879	Sequence 22879, A
43	40	43.5	614	4	US-09-252-991A-29695	Sequence 29695, A
44	40	43.5	676	2	US-08-398-590A-40	Sequence 40, Appl
45	40	43.5	676	3	US-08-894-997-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-22901
; Sequence 22901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22901
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22901

Query Match 51.1%; Score 47; DB 4; Length 629;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 QESTGRSGRSGXSGS 19
Db 272 RQSHRGRAAGRAHRS 288

RESULT 2
US-09-252-991A-27836
; Sequence 27836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27836
; LENGTH: 157
; TYPE: PRT

```
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27836

Query Match          47.8%; Score 44; DB 4; Length 157;
Best Local Similarity 66.7%; Pred. No. 9.5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 TGGRRGRSGXS 17
   |||||
Db 70 SRGRGRSGSA 81

RESULT 3
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Cardamine oligosperma
; US-09-513-057C-13

Query Match          47.8%; Score 44; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 42;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGXS 19
   |||||
Db 573 SRQVTSASASGRGSGS 591

RESULT 4
US-09-252-991A-27107
; Sequence 27107, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27107
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-27107

Query Match          46.7%; Score 43; DB 4; Length 159;
Best Local Similarity 57.9%; Pred. No. 14;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTRGRSGRSGXS 19
   |||||
Db 5 SASSTRGRSAMRSTASS 23

RESULT 5
US-09-252-991A-25519
; Sequence 25519, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25519
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25519

Query Match          46.7%; Score 43; DB 4; Length 581;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQSTRGRSGRSGXS 19
   |||||
Db 178 HRADSRGRSRERRGAPRS 195

RESULT 6
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-984-890-4

Query Match          46.7%; Score 43; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGRSGRSGXS 19
   |||||
Db 562 SPQSHSGRGRGASGS 576

RESULT 7
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
```



```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-984-890-2

Query Match
Best Local Similarity 46.7%; Score 43; DB 4; Length 724;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGSRGRSGXSGS 19
Db 564 SPGSHSQRRGASGS 578

RESULT 8
US-09-523-849-36
/ Sequence 36; Application US/09523849
/ Patent No. 6458561
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Molteni, Angela
/ APPLICANT: Magnaghi, Paola
/ APPLICANT: Bosotti, Roberta
/ APPLICANT: Scaccheri, Emanuela
/ APPLICANT: Issacchi, Antonella
/ APPLICANT: Hodgson, Dave
/ TITLE OF INVENTION: HUMAN NIM1 KINASE
/ FILE REFERENCE: PC-0009 US
/ CURRENT APPLICATION NUMBER: US/09/523,849
/ CURRENT FILING DATE: 2000-03-13
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PERL Program
/ SEQ ID NO 36
/ LENGTH: 745
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Genbank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match
Best Local Similarity 46.7%; Score 43; DB 4; Length 745;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STRGSRGRSGXSGS 19
Db 585 SPGSHSQRRGASGS 599

RESULT 9
US-08-354-973-1
/ Sequence 1; Application US/08354973
/ Patent No. 5641666
/ GENERAL INFORMATION:
/ APPLICANT: Vahlensieck, Hans-Friedrich
/ APPLICANT: Hinnen, Albert
/ TITLE OF INVENTION: Fungi Resistant to Soraphen A
/ NUMBER OF SEQUENCES: 1
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ciba-Geigy Corporation
/ STREET: Patent Dept., 520 White Plains Rd., POB 2005
/ CITY: Tarrytown
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10591-9005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/354,973
/ FILING DATE: 13-Dec-1994
```

```
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Pace, Gary
/ REGISTRATION NUMBER: 40,403
/ REFERENCE/DOCKET NUMBER: PF/5-19802/A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8582
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2237 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Protein
/ HYPOTHEICAL: NO
US-08-354-973-1

Query Match
Best Local Similarity 46.7%; Score 43; DB 1; Length 2237;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTGRSRGRSGXSGS 19
Db 1200 HOSSNGCAPDRSGSSAS 1217

RESULT 10
US-09-252-991A-19488
/ Sequence 19488; Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 19488
/ LENGTH: 398
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19488

Query Match
Best Local Similarity 46.2%; Score 42.5; DB 4; Length 398;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 2 HOESTGRSRGRSG 15
Db 326 HRESRRGRSGGRKG 340

RESULT 11
US-09-252-991A-32394
/ Sequence 32394; Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
```

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32394
LENGTH: 105
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32394

Query Match 45.7%; Score 42; DB 4; Length 105;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 TRGSRGRSGXSG 18
DB 65 TRGRCACRCATG 77

RESULT 12
US-09-252-991A-27336
Sequence 27336, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27336
LENGTH: 349
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27336

Query Match 45.7%; Score 42; DB 4; Length 349;
Best Local Similarity 52.6%; Pred. No. 45;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTGRSGRSGXSGS 19
DB 309 SRPASGRGRCRCSAGAGS 327

RESULT 13
US-09-252-991A-21124
Sequence 21124, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21124
LENGTH: 523
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21124

Query Match 45.7%; Score 42; DB 4; Length 523;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQESTGRSGRSGXSG 18
DB 85 SRPAATRGRRRRRGLG 102

RESULT 14
US-08-724-548-54
Sequence 54, Application US/08724548
Patent No. 5830637
GENERAL INFORMATION:
APPLICANT: Frank, Ronald
APPLICANT: Guler, Sinan
TITLE OF INVENTION: PROCESS FOR THE RAPID SYNTHESIS OF
TITLE OF INVENTION: SUPPORT-BOUND OR FREE PEPTIDES OR
TITLE OF INVENTION: OLIGONUCLEOTIDES, A FLAT MATERIAL,
TITLE OF INVENTION: PREPARED THEREWITH AND ITS USE AND
TITLE OF INVENTION: A DEVICE FOR CARRYING OUT THE
PROCESS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSER: Joseph T. Eisele
ADDRESSER: Kane, Dalsimer, Sullivan, Kurucz,
ADDRESSER: Levy, Eisele and Richard
STREET: 711 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017-4059
COMPUTER READABLE FORM:
MEDIUM TYPE: 3-1/2" DISKETTE
COMPUTER: IBM-XT COMPATIBLE
OPERATING SYSTEM: DOS 3.3:
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,548
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,674
FILING DATE: 03/22/93
ATTORNEY/AGENT INFORMATION:
NAME: EISELE, JOSEPH T.
REGISTRATION NUMBER: 25,331
REFERENCE/DOCKET NUMBER: 2727-68 PCT
TELEPHONE: (212) 687-6000
TELEFAX: (212) 682-3485
TELEX: (212) 426767
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 residues
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: Peptide
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
US-08-724-548-54

Query Match 44.6%; Score 41; DB 2; Length 70;
Best Local Similarity 61.5%; Pred. No. 12;

Tue Sep 28 07:11:33 2004

us-09-308-150-5.open.ra1

Page 5

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 7 RGRSRGRSGXSGS 19
Db 7 RGRSRGRSGXSGS 19

Query Match 44.6%; Score 41; DB 3; Length 70;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 7 RGRSRGRSGXSGS 19
Db 7 RGRSRGRSGXSGS 19

Search completed: September 28, 2004, 06:26:43
Job time : 14.4479 secs

RESULT 15
US-07-978-674B-54
; Sequence 54, Application US/07978674B
; Patent No. 6040423
; GENERAL INFORMATION:
; APPLICANT: Frank, Ronald
; APPLICANT: Guier, Simon
; TITLE OF INVENTION: PROCESS FOR THE RAPID SYNTHESIS OF
; TITLE OF INVENTION: SUPPORT-BOUND OR FREE PEPTIDES OR
; TITLE OF INVENTION: OLIGONUCLEOTIDES, A FLAT MATERIAL
; TITLE OF INVENTION: PREPARED THEREWITH AND ITS USE AND
; TITLE OF INVENTION: A DEVICE FOR CARRYING OUT THE
; TITLE OF INVENTION: PROCESS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Joseph T. Eisele
; ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,
; ADDRESSEE: Levy, Eisele and Richard
; STREET: 711 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10017-4059
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/2" DISKETTE
; COMPUTER: IBM-XT COMPATIBLE
; OPERATING SYSTEM: DOS 3.3;
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/978,674B
; FILING DATE: 03/22/93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: EISELE, JOSEPH T.
; REGISTRATION NUMBER: 25,331
; REFERENCE/DOCKET NUMBER: 2727-68 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 687-6000
; TELEFAX: (212) 682-3485
; TELEX: (212) 426767
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 70 residues
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: Peptide
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLER:
; IMMEDIATE SOURCE:
US-07-978-674B-54

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds

(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-6
Perfect score: 88
Sequence: 1 SHOESTXGSRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	2248	2	A35938
2	78	88.6	416	2	A32947
3	69	78.4	591	2	A45135
4	52	59.1	506	1	W2ML47
5	49	55.7	822	2	T51049
6	47	53.4	399	2	T46259
7	44	50.0	800	2	T02852
8	44	50.0	827	2	T13468
9	44	50.0	842	1	JDVLVS
10	44	50.0	843	1	JDVLVR
11	43	48.9	306	2	T21220
12	43	48.9	694	2	S71786
13	43	48.9	849	2	A96592
14	42	47.7	136	2	T35632
15	42	47.7	315	2	S53589
16	42	47.7	471	2	T33997
17	42	47.7	745	2	G01025
18	42	47.7	774	2	I48609
19	42	47.7	889	2	T33422
20	42	47.7	1337	2	T30291
21	42	47.7	1804	2	T34518
22	41.5	47.2	1895	2	T06609
23	41	46.6	825	2	GC4163
24	41	46.6	836	2	GC4727
25	41	46.6	843	1	T43491
26	41	46.6	1015	2	T42013
27	41	46.6	2233	2	S63347
28	40	45.5	163	2	T23076
29	40	45.5	167	2	S05545

30	40	45.5	168	2	A39316	RAB-17 protein - m
31	40	45.5	168	2	S08633	RAB-17 protein - m
32	40	45.5	197	2	T18918	hypothetical prote
33	40	45.5	203	2	C87801	protein C10G11.9 l
34	40	45.5	261	2	S33356	hypothetical prote
35	40	45.5	269	2	B38095	T-cell-specific tr
36	40	45.5	272	2	A38900	T-cell-specific tr
37	40	45.5	303	2	JH0401	130k paracrystall
38	40	45.5	312	2	A31846	hypothetical prote
39	40	45.5	508	2	T22440	T20H2.2 protein -
40	40	45.5	538	2	H65335	nucleolin homolog
41	40	45.5	611	2	T06458	probable ATP-depen
42	40	45.5	649	2	G71283	hypothetical prote
43	40	45.5	692	2	H71426	hypothetical prote
44	40	45.5	695	2	G85135	hypothetical prote
45	40	45.5	708	2	T25725	hypothetical prote

ALIGNMENTS

RESULT 1
A35938
Profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence__revision 02-Jul-1996 #text__change 29-Sep-1999
C/Accession: A35938
R/Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinhert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; PMID:91064347; PMID:2248557
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:J02929
C/Genetics:
A/Gene: GDB:RUG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermidis; polymorphism; tandem repeat
F:246-569/Region: profilaggrin repeat
F:570-893/Region: profilaggrin repeat
F:1074-1397/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 95.5%; Score 84; DB 2; Length 2248;
Best Local Similarity 89.5%; Pred. No. 2.8e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGSRGSRGSGS 19
Db 551 SHOESTRGSRGSRGSGS 569

RESULT 2
A32947
Filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence__revision 04-Sep-1992 #text__change 29-Sep-1999
R/McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of tr
A/Reference number: A32947; PMID:8926901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-references: GB:M2435; NID:g182604; PTDN:AA52454.1; PID:g182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:

A;Gene: GDB:FLG
 A;Cross-references: GDB:119912; OMIM:135940
 A;Map position: 1q21-1q21
 C;Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C;Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 88.6%; Score 78; DB 2; Length 416;
 Best Local Similarity 84.2%; Pred. No. 4.9e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGXSRRSGRSGS 19
 |||||
 DB 7 SHQESTRGRSAGRSGRSGS 25

RESULT 3

A45135
 A;Protein: Profilaggrin - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C;Accession: A45135
 R;Presland, R.B.; Haydock, P.V.; Fleckman, P.; Niyonsukxiri, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992
 A;Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A;Reference number: A45135; PMID:93054736; PMID:1429717
 A;Accession: A45135
 A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA
 A;Residues: 1-591 <PR>
 A;Cross-references: GB:L01089; GB:M90967; NID:G190408; PID:AAA60177.1; PID:G553621
 A;Note: sequence extracted from NCBI backbone (NCBIP:118773)
 C;Genetics:

A;Gene: GDB:FLG

A;Cross-references: GDB:119912; OMIM:135940

A;Map position: 1q21-1q21
 C;Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C;Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F;49-81/Dominant: calmodulin repeat homology <EP2>

Query Match 78.4%; Score 69; DB 2; Length 591;
 Best Local Similarity 78.9%; Pred. No. 0.0019;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHQESTXGXSRRSGRSGS 19
 |||||
 DB 449 SHQESTRGRSAGRSGRSGS 467

RESULT 4

W2W147
 E2 protein - human papillomavirus type 47
 C;Species: human papillomavirus type 47
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C;Accession: D35324
 R;Kiyono, T.; Adachi, A.; Ishibashi, M.
 Virology 177, 401-405, 1990
 A;Title: Genome organization and taxonomic position of human papillomavirus type 47 infe

A;Reference number: A35324; PMID:90281611; PMID:2162112
 A;Accession: D35324
 A;Status: translation not shown

A;Molecule type: DNA
 A;Residues: 1-506 <KIT>

A;Cross-references: GB:M23305; NID:G333062; PID:AAA46979.1; PID:G333067
 C;Superfamily: papillomavirus E2 protein
 C;Keywords: DNA binding; early protein; transcription regulation

Query Match 59.1%; Score 52; DB 1; Length 506;
 Best Local Similarity 52.6%; Pred. No. 0.89;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTXGXSRRSGRSGS 19
 |||||

DB 342 SRGNTTGRGRGRGRG 360

RESULT 5

T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N;Alternate names: protein B12P1.10
 C;Species: Neurospora crassa
 C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C;Accession: T51049
 R;Schulte, U.; Algn, V.; Hohlseisel, U.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000

A;Reference number: Z25286
 A;Accession: T51049
 A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-822 <SCH>

A;Cross-references: EMBL:AL390091; GSPDB:GNO0116; NCSP:B12P1.10

A;Experimental source: BAC clone B12P1; strain OR74A

C;Genetics:

A;Gene: NCSP:B12P1.10

A;Map position: 6

A;Inserts: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 4.3;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGXSRRSGRSGS 19
 |||||
 DB 414 RESASGRTRGRGRGRGT 430

RESULT 6
 T46259
 hypothetical protein DKFZp761E0323.1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
 C;Accession: T46259

R;Boecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23032

A;Accession: T46259

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-399 <AAA>

A;Cross-references: EMBL:AL137423

A;Experimental source: adult amygdala; clone DKFZp761E0323

C;Genetics:

A;Note: DKFZp761E0323.1

Query Match 53.4%; Score 47; DB 2; Length 399;
 Best Local Similarity 55.6%; Pred. No. 4.5;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTXGXSRRSGRSGS 18
 |||||
 DB 336 SRSRGRSGRSGRSGS 353

RESULT 7
 T02852
 probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)

C;Species: Leishmania major
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
 C;Accession: T02852; H81462
 R;Wyley, P.J.

submitted to the EMBL Data Library, May 1998

A;Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.

A;Reference number: Z14740

A;Accession: T02852

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A:Residues: 1-800 <MYL>
 A:Cross-references: EMBL:AE001274; NID:G3264850; PID:G2266920
 R:Myler, P.J.; Audleman, L.J.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
 A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
 A:Reference number: AB1455; MUID:99178987; PMID:10077609
 A:Accession: AB1462
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-800 <PYL>
 A:Cross-references: GB:AE001274; NID:G3264850; PID:ACC24675.1; PID:G2266920; GSPDB:GN00
 A:Experimental source: strain MHW/IL/81/Friedlin
 C:Genetics:
 A:Gene: L1439.4
 A:Map position: 1

Query Match 50.0%; Score 44; DB 2; Length 800;
 Best Local Similarity 50.0%; Pred. No. 27;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 HOESTXGSRGRSGSGS 19
 ||:|||||
 Db 429 HRDGVRLSTVSGRSGS 446

RESULT 8
 T13468
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)
 C:Species: hepatitis B virus, HBV
 A:Variety: isolate 09D09HCC
 C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
 C:Accession: T13468
 R:Takehashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishihiro, S.
 Arch. Virol. 143, 2313-2326, 1998
 A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
 A:Reference number: Z17684; MUID:99129050; PMID:9930189
 A:Accession: T13468
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-827 <TAK>
 A:Cross-references: EMBL:AB014368; NID:G3551304; PID:BA32863.2; PID:G6116707
 A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
 C:Genetics:
 A:Gene: P
 A:Introns: 303/3
 A:Superfamily: hepatitis virus DNA-directed DNA polymerase
 C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 2; Length 827;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILAKSGRSGS 242

RESULT 9
 JDUVLV
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
 C:Species: hepatitis B virus, HBV
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 25-Oct-1996
 C:Accession: S04568
 R:Pho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
 Nucleic Acids Res. 17, 2124, 1989
 A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp
 A:Reference number: S04568; MUID:89183619; PMID:2928116
 A:Accession: S04568
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-842 <RHO>
 A:Cross-references: EMBL:X14193
 C:Superfamily: hepatitis virus DNA-directed DNA polymerase

C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 1; Length 842;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILAKSGRSGS 242

RESULT 10
 JDUVLV
 DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
 C:Species: hepatitis B virus, HBV
 C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 25-Oct-1996
 C:Accession: A00704
 R:Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
 Nucleic Acids Res. 11, 1747-1757, 1983
 A:Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
 A:Reference number: A93460; MUID:83168919; PMID:6300776
 A:Accession: A00704
 A:Molecule type: DNA
 A:Residues: 1-843 <ONO>
 A:Superfamily: hepatitis virus DNA-directed DNA polymerase
 C:Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 50.0%; Score 44; DB 1; Length 843;
 Best Local Similarity 52.9%; Pred. No. 28;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGRSGSGS 19
 ||:|||||
 Db 226 QPQGSILAKSGRSGS 242

RESULT 11
 T21220
 hypothetical protein F21H7.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T21220
 R:Gardner, A.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19393
 A:Accession: T21220
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-306 <ML>
 A:Cross-references: EMBL:Z93379; PID:CA807588.1; GSPDB:GN00023; CESP:F21H7.5
 A:Experimental source: clone F21H7
 C:Genetics:
 A:Gene: CESP:F21H7.5
 A:Map position: 5
 A:Introns: 28/3; 146/3; 232/3
 A:Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHQSTXGSRGRSGSGS 18
 ||:|||||
 Db 48 SHRSKTKRSGRSGSGS 65

RESULT 12
 S71786
 wingless receptor precursor dfx2 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 C:Accession: S71786; S78444
 R:Bhanoc, P.; Brink, M.; Harryman Samos, C.; Heileh, J.C.; Wang, Y.; Macke, J.P.; Andrew,

Nature 382, 225-230, 1996
 A>Title: A new member of the frizzled family from *Drosophila* functions as a wingless rec
 A:Reference number: S71786; PMID:96353971; PMID:8717036
 A:Accession: S71786
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1694 <BHA>
 A:Cross-references: EMBL:U65589
 A>Note: mRNA was also sequenced
 R:Bhanot, P.; Wang, Y.; Nathans, J.
 submitted to the EMBL Data Library, July 1996
 A:Reference number: S78444
 A:Accession: S78444
 A:Molecule type: DNA
 A:Residues: 1416, 'T', 418-694 <BHM>
 A:Cross-references: EMBL:U65589; NID:g1518050; PIDN:AAC47273.1; PID:g1518051
 A:Gene: dfz2
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0016797
 C:Superfamily: fruit fly frizzled protein
 C:Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 Oy 1 SHOESTXGSGRSGRSGS 19
 Db 182 STTASGSGSGSGSGSGS 200

RESULT 13
 A96592
 Hypothetical protein F14C21.55 [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96592
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Hansen, N.F.; Hughes, B.; Hultz, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzall,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; PMID:21016719; PMID:11130712
 A:Accession: A96592
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-849 <STO>
 A:Cross-references: GB:AE005173; NID:g11055759; PIDN:AG28231.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F14C21.55
 A:Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 5 STXGSGRSGRSGS 19
 Db 736 SSANRSGRSGRSGS 750

RESULT 14
 T35632
 Probable transposase - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
 C:Accession: T35632
 R:Seeger, K.U.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999
 A:Reference number: Z21584
 A:Accession: T35632
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-136 <SEE>
 A:Cross-references: EMBL:AU079356; PIDN:CAB45627.1; GSPDB:GN00070; SCOEDB:SC6G9.36C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC6G9.36C
 C:Superfamily: *Streptomyces coelicolor* probable transposase SC6G9.36C

Query Match 47.7%; Score 42; DB 2; Length 136;
 Best Local Similarity 41.2%; Pred. No. 9.8;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 SHOESTXGSGRSGRSGS 17
 Db 111 AHQAAAGARRKGRGRRGA 127

RESULT 15
 S53589
 SOL2 protein - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein YCR073w-a; protein YCRX13w
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 19-May-1995 #sequence_revision 17-Nov-1995 #text_change 19-Apr-2002
 C:Accession: S53589; S70385; S53593
 R:Jimenez, A.
 submitted to the EMBL Data Library, December 1992
 A:Reference number: S53589
 A:Accession: S53589
 A:Molecule type: DNA
 A:Residues: 1-315 <JIM>
 A:Cross-references: EMBL:X59720; NID:g1907116; PID:g1907213
 R:Shen, W.C.; Stanford, D.R.; Hopper, A.K.
 Genetics 143, 699-712, 1996
 A>Title: *Los1p*, involved in yeast pre-tRNA splicing, positively regulates members of the
 A:Reference number: S70383; PMID:96363906; PMID:8725220
 A:Accession: S70385
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-170, 'A', 172-315 <SHE>
 A:Cross-references: EMBL:U46559; NID:g1184940; PIDN:AA849322.1; PID:g1184941
 C:Genetics:
 A:Gene: SGD:SOL2
 A:Cross-references: MIPS:YCR073w-a; SGD:S0000718
 A:Map position: 3R
 C:Function:
 A:Description: multicopy suppressor of *los1-1*

Query Match 47.7%; Score 42; DB 2; Length 315;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Oy 4 ESTXGSGRSGRSGS 19
 Db 60 KSTASAGSGSGSGS 75

Search completed: September 28, 2004, 06:15:01
 Job time : 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-6

Perfect score: 88
Sequence: 1 SHQSTXGXSRRGSRGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	88.6	416	1	FILA_HUMAN
2	52	59.1	416	1	VE2_HPV47
3	44	50.0	730	1	DPOL_HPBV4
4	44	50.0	842	1	DPOL_HPBVM
5	44	50.0	843	1	DPOL_HPBVR
6	43	48.9	694	1	FRZ2_DROME
7	42	47.7	315	1	SOL2_YEAST
8	42	47.7	774	1	MRX2_MOUSE
9	42	47.7	1337	1	DEXT_STRDO
10	41.5	47.2	1895	1	WR19_ARATH
11	41	46.6	825	1	5B5_RAT
12	41	46.6	1015	1	FRQ_CRESP
13	41	46.6	2233	1	COX6_YEAST
14	40.5	46.0	379	1	ROA3_MOUSE
15	40	45.5	168	1	PDH1_MAIZE
16	40	45.5	380	1	PEXD_PICPA
17	40	45.5	419	1	TCF7_MOUSE
18	40	45.5	660	1	DDX7_HUMAN
19	40	45.5	778	1	SEA2_MYCLE
20	40	45.5	808	1	SEA2_MYCTU
21	40	45.5	955	1	CLS2_HUMAN
22	39.5	44.9	135	1	RBP1_DROME
23	39	44.3	155	1	RR7_PSIINU
24	39	44.3	439	1	AP50_DICDI
25	39	44.3	508	1	V202_HUMAN
26	39	44.3	686	1	BDPI_YEAST
27	38	43.2	325	1	Y42A_RHIST
28	38	43.2	373	1	YF04_YEAST
29	38	43.2	502	1	VE2_HPV25
30	38	43.2	643	1	GAG_HPV31
31	38	43.2	645	1	RIPR_MOUSE
32	38	43.2	651	1	SEC9_YEAST
33	38	43.2	824	1	TGM1_RAT

34	38	43.2	846	1	IF2_PSEPK	Q884V7 pseudomonas
35	38	43.2	965	1	CLS2_RAT	Q8vd41 rattus norv
36	38	43.2	966	1	CLS2_MOUSE	Q9et65 mus musculu
37	38	43.2	978	1	PEX6_RAT	P54777 rattus norv
38	38	43.2	980	1	PEX6_HUMAN	Q13608 homo sapien
39	38	43.2	1094	1	ABBI_HUMAN	Q00203 homo sapien
40	38	43.2	1232	1	Y005_CAEEL	P34643 caenorhabd1
41	38	43.2	1383	1	PRAX_RAT	Q63445 rattus norv
42	38	43.2	1733	1	VNUA_PRYKA	P33465 pseudorabie
43	37.5	42.6	832	1	DPOL_HPBVL	P12900 hepatitis b
44	37	42.0	123	1	LSM4_CAEEL	Q19952 caenorhabd1
45	37	42.0	136	1	SR19_ORYSA	P49964 oryza sativ

ALIGNMENTS

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
ID FILA_HUMAN
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69296901; Pubmed=2740331;
RA McIntirey-Grant U.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Leasin S.R., Steinert P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RL [2]
RP CITRULLINATION.
RX MEDLINE=96374388; Pubmed=8780679;
RA Senshu T., Kan S., Ogawa H., Manabe M., Aaga H.;
RT "Preferential determination of keratin K1 and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -!- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -!- PTM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -!- PTM: Undergoes determination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M24355; AAA52454.1; -
DR PIR: A32947; A32947.
DR Genew; HGNC:3748; FLG.
DR MIM: 135940; -
DR GO: GO:0005882; C:intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007275; P:development; NAS.
DR InterPro: IPR003303; Filaggrin.

DR Pfam: PF03516; Filaggrin; 2.
 DR PRINTS: PR00487; FILAGGRIN.
 KM Phosphorylation; Citrullination; Developmental protein.
 FT NON_TER
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 88.6%; Score 78; DB 1; Length 416;
 Best Local Similarity 84.2%; Pred. No. 1,4e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQESTXGKSRGSRGSGS 19
 |||||
 Db 7 SHQESTRGSRGSRGSGS 25

RESULT 2
 ID_VB2_HPVA7 STANDARD; PRT; 506 AA.

AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.

OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC Papillomavirus.
 OX NCBI_TaxID=10594;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=90281611; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.
 RT "Genome organization and taxonomic position of human papillomavirus
 type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).

CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCCNNNNNGGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING OF E2RE'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.

CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M32305; AAA46979.1; -.
 DR PIR: D35324; M2ML47.
 DR HSSP: P03122; 2BOP.
 DR InterPro: IPR000427; E2_C.
 DR InterPro: IPR001866; E2_N.
 DR InterPro: IPR009021; Viral_DNA_bd.
 DR Pfam: PF00511; E2_C; 1.
 DR Pfam: PF00508; E2_N; 1.
 DR ProDom: PD000672; E2_C; 1.
 DR ProDom: PD000678; E2_N; 1.
 KM Early protein; Transcription regulation; Activator; DNA-binding;
 KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F9BFB725065 CRC64;

Query Match 59.1%; Score 52; DB 1; Length 506;
 Best Local Similarity 52.6%; Pred. No. 0.32;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQESTXGKSRGSRGSGS 19

Db 342 SRENTRGRRGRRGRRGS 360

RESULT 3
 ID_DPOL_HPBVA STANDARD; PRT; 730 AA.

AC P12933;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
 DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]
 GN P.

OS Hepatitis B virus (subtype adr4).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 CC NCBI_TaxID=10409;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=83246570; PubMed=6306594;
 RA Fujiyama A., Miyahara A., Nozaki C., Yoneyama T., Ohtomo N.,
 RA Matsubara K.;
 RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
 RT adr.";

RL Nucleic Acids Res. 11:4601-4610(1983).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA} (N).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphomonoester.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X01587; CA25746.1; -.
 DR InterPro: IPR001462; DNAPol_viral_C.
 DR InterPro: IPR00201; DNAPol_viral_N.
 DR InterPro: IPR000477; RYase.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; rvt; 2.
 DR ProDom: PD000814; DNAPol_viral_C; 1.
 KM Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
 KM Hydrolyase; Nuclease; Endonuclease; DNA replication; DNA-binding.
 SQ SEQUENCE 730 AA; 82207 MW; 898A8BBB50D496F2 CRC64;

Query Match 50.0%; Score 44; DB 1; Length 730;
 Best Local Similarity 52.9%; Pred. No. 9.8;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
 |||||
 Db 226 QPQGSRLRGSRGSGS 242

RESULT 4
 ID_DPOL_HPBVM STANDARD; PRT; 842 AA.

AC P31870; Q67861; Q67866;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
 DE directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)]
 GN P.

OS Hepatitis B virus (subtype adr / mutant).
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 CC NCBI_TaxID=31512;
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=89183619; PubMed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
virus subtype adr.";
RL Nucleic Acids Res. 17:2124-2124(1989).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14193; CAA32399.1; ALT_TERM.
CC EMBL; X14193; CAA32405.1; ALT_TERM.
CC PIR; S04568; JDVLVS.
CC InterPro; IPR001462; DNAPol_viral_C.
CC InterPro; IPR000201; DNAPol_viral_N.
CC InterPro; IPR000477; RTase.
CC Pfam; PF00336; DNA_pol_viral_C; 1.
CC Pfam; PF00242; DNA_pol_viral_N; 1.
CC Pfam; PF00078; rvt; 2.
CC ProDom; PD000814; DNAPol_viral_C; 1.
CC Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
CC Hydroxylase; Nuclease; Endonuclease; DNA replication; DNA-binding.
CC SEQUENCE 842 AA; 94545 MW; 137FE37246BA5C6C CRC64;

Query Match 50.0%; Score 44; DB 1; Length 842;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGKSGRSGRSGS 19
Db 226 QPOGSLARGKSGRSGS 242

RESULT 5
ID DPOL_HPBVR STANDARD; PRT; 843 AA.
AC P03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
OS Hepatitis B virus (subtype adr).
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=106820;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168919; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adw.";
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA} (N).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00867; NOT_ANNOTATED_CDS.
CC PIR; A00704; JDVLVR.
CC InterPro; IPR001462; DNAPol_viral_C.
CC InterPro; IPR000201; DNAPol_viral_N.
CC InterPro; IPR000477; RTase.
CC Pfam; PF00336; DNA_pol_viral_C; 1.
CC Pfam; PF00242; DNA_pol_viral_N; 1.
CC Pfam; PF00078; rvt; 2.
CC ProDom; PD000814; DNAPol_viral_C; 1.
CC Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
CC Hydroxylase; Nuclease; Endonuclease; DNA replication; DNA-binding.
CC SEQUENCE 843 AA; 94400 MW; A6B2D49083C4E8B CRC64;

Query Match 50.0%; Score 44; DB 1; Length 843;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGKSGRSGRSGS 19
Db 226 QPOGSLARGKSGRSGS 242

RESULT 6
ID FRZ2_DROME STANDARD; PRT; 694 AA.
AC Q9VFX3; Q94916; Q9VFX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Frizzled protein 2 precursor (Frizzled-2) (dfrz2).
DE FZ2 OR CG9739/CG14083.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CC NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A. AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
RX MEDLINE=96353971; PubMed=8717036;
RA Bhancoc P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
RA Andrew D., Nathans J., Nuse R.;
RT "A new member of the frizzled family from Drosophila functions as a
RT Wingless receptor";
RL Nature 382:225-230(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Berens P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butts K.C., Cawley S.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Chew J.S., Dahlke C., Davaport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idagwam C.,
RA Jajati M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacted J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slater E., Spading A.C., Stapleton M., Strong R., Sun E.,
RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC
CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
CC are coupled to the beta-catenin canonical signaling pathway, which
CC leads to the activation of dishevelled proteins. Inhibition of
CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
CC of Wnt target genes. A second signaling pathway involving PKC and
CC calcium fluxes has been seen for some family members, but it is
CC not yet clear if it represents a distinct pathway or if it can be
CC integrated in the canonical pathway, as PKC seems to be required
CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
CC to involve interactions with G-proteins. Required to coordinate
CC the cytoskeletons of epidermal cells to produce a parallel array
CC of cuticular hairs and bristles.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
CC between 15 and 70 per cent of egg length, including the
CC invaginating cells of the ventral furrow. Stripe pattern is
CC emerging by early stage 8. From stage 9 and continuing throughout
CC embryogenesis, expression is seen in the developing CNS. At stage
CC 10, expressed in 15 stripes in the presumptive head and trunk
CC regions, in the posterior midgut primordium, in a subset of cells
CC of anterior midgut invagination and in the procephalic lobe. At
CC stage 12, expression declines in epidermis and increases in the
CC midgut and visceral mesoderm. At stage 17, only expressed in the
CC CNS, hindgut and dorsal vessel.
CC
CC -1- DOMAIN: Lys-Thr-X-X-Tyr motif is involved in the activation of
CC the Wnt/beta-catenin signaling pathway (By similarity).
CC
CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
CC
CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
CC family.
CC
CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: U65589; AAC47273.1; -;
CC EMBL: AE003518; AAF49185.2; -;
CC PIR: S71786;
CC TrEMBL: F8B90016797; f22;
CC
CC DR GO: GO:0016021; C: integral to membrane; NMS.
CC DR GO: GO:0017147; F: Wnt-protein binding; IDA.
CC DR GO: GO:0006928; P: cell motility; IMP.
CC DR GO: GO:0007163; P: establishment and/or maintenance of cell po. .; ISS.
CC DR GO: GO:0005855; P: female gonad development; IMP.
CC DR GO: GO:0016055; P: Wnt receptor signaling pathway; IDA.
CC
CC DR InterPro: IPR000539; Frizzled.
CC DR InterPro: IPR000024; Fz domain.
CC DR InterPro: IPR000832; GPCR_sectretin.
CC DR Pfam: PF01534; Frizzled_1.
CC DR Pfam: PF01392; Fz_1.
CC DR PRINTS: PRO0489; FRIZZLED.
CC SMART: SM00063; FRI; 1.
CC DR PROSITE: PSS0038; FZ; 1.
CC DR PROSITE: PSS0261; G_PROTEIN_RECEP_F2_4; 1.

KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
KM Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 694
FT DOMAIN 23 315
FT DOMAIN 316 336
FT DOMAIN 337 352
FT DOMAIN 353 373
FT DOMAIN 374 397
FT TRANSM 398 418
FT TRANSM 419 439
FT TRANSM 440 460
FT TRANSM 461 482
FT TRANSM 483 503
FT TRANSM 504 534
FT TRANSM 535 555
FT TRANSM 556 584
FT TRANSM 585 605
FT TRANSM 606 694
FT DOMAIN 187 225
FT DOMAIN 59 180
FT SITE 608 613
FT SITE 692 694
FT CARBOHYD 78 78
FT CARBOHYD 288 288
FT CONFLICT 55 55
FT CONFLICT 417 417
SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAFB096 CRC64;

Qy 1 SHOSTYXSGRGRSGSGS 19
Db 162 SYTAGSGSGSGSGSGS 200

Query Match 48.9%; Score 43; DB 1; Length 694;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

RESULT 7
SOL2_YEAST STANDARD; PRT; 315 AA.
ID SOL2_YEAST
AC P37262;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (P6GL).
GN SOL2 OR YCRX13W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96363906; PubMed=8725220;
RA Shen W.C., Stanford D.R., Hopper A.K.;
RT "Loss of a yeast pre-tRNA splicing, positively regulates
RT members of the SOL gene family.";
RL Genetics 143:699-712(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Ballesca J.P.G., Franco L., Hoenicka U., Jimenez A., Remacha M.,
RA Sanz E.;
RT Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION, AND SIMILARITY.
RX MEDLINE=94147995; PubMed=8313894;
RA Kocin E.V., Bork P., Sander C.;
RT "Yeast chromosome III: new gene functions.";
RL EMBO J. 13:493-503(1994).
CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
CC phosphogluconate (By similarity).
CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC phospho-D-gluconate.

```

CC -1- PATHWAY: Pentose phosphate pathway; second step.
CC -1- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate
CC isomerase family. 6-phosphoglucosaminolactonase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U46559; AA349322.1; -.
DR EMBL: X59720; CAA42272.1; -.
DR PIR: S53589; S53589.
DR GeneBank: 139004; -.
DR SGD: S0000718; SOL2.
DR GO: GO:0008033; P:RNA processing; IGI.
DR InterPro: IPR006148; Gluc gal isom.
DR InterPro: IPR005900; Phosphoglucolac.
DR Pfam: PF01182; Glucosamine iso; 1.
DR TIGRfam: TIGR01198; pgl_1.
DR HydroLase.
KM CONFLICT 171 171 A -> V (IN REF. 2).
SQ SEQUENCE 315 AA; 34501 MW; 7F0AAD76574AB276 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 315;
Best Local Similarity 50.0%; Pred. No. 8.4;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 ESTXGSRGSRGSGS 19
Db 60 KSTASAAEGKSGSGS 75

RESULT 9
MRK2_MOUSE STANDARD; PRT; 774 AA.
ID MRK2_MOUSE
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
GN MARK2 OR EMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93364122; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19.";
RL Mamm. Genome 4:401-403 (1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: X70764; CA550040.1; -.
DR PIR: I48609; I48609.
DR HSSP: Q63450; I406.
DR MGD; MGI:99638; Mark2.

```

```

DR InterPro: IPR001772; Kinase Cterm.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF02149; KAI; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase: Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 53 304
FT NP_BIND 59 67 ATP (BY SIMILARITY).
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BFED7BF443483A CRC64;

Query Match 47.7%; Score 42; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGSRGSRGSGS 19
Db 614 SPESHSGRGRGSGS 628

RESULT 9
DXT STRDO STANDARD; PRT; 1337 AA.
ID DXT STRDO
AC P39653;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
DE glucanohydrolase).
GN DEX.
OS Streptococcus downei (Streptococcus sobrinus).
OC Plasmid pXA902.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
RC STRAIN=6715 / UAB66;
RX MEDLINE=94292401; PubMed=8021165;
RA Wanda S.-Y., Curtiss R. III;
RT "Purification and characterization of Streptococcus sobrinus
RT dextranase produced in recombinant Escherichia coli and sequence
RT analysis of the dextranase gene.";
RL J. Bacteriol. 176:3839-3850 (1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
CC PELICLE-COATED TOOTH SURFACE.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
CC linkages in dextran.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
CC and at 39 degrees Celsius.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----

```

```

DR EMBL, M66978; AAA21772.1; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfam; TIGR0167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Hydroxylase; Glycosylase; Cell wall; Peptidoglycan-anchor; Repeat;
KW Signal; Plasmaid.
FT CHAIN 1 30
FT SIGNAL 31 1308
FT PROPEP 1309 1337
FT SITE 1305 1309
FT MOD_RES 1308 1308
SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SHOEKTXGSGRSGRS 17
Db 1154 ANQSTKGSADSGSKS 1170

RESULT 10
WR19 ARATH STANDARD; PRT; 1895 AA.
AC 093267;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Probable WRKY transcription factor 19 (WRKY DNA-binding protein 19).
GN WRKY19 OR At4G12020 OR Flj013.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid 1; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10615198;
RA Meyer K.F.X., Schnell R., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansoorge W., Brandt P., Griwell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Maehre R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohelsel J., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schuren J., Grymoprez B., Chuang Y.-J., Vandebussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Wiltzengeser T., Bothe G., Rampsberger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., Van Staveren M., Dirke W.,
RA Moolman P., Klein lankhorst R., Rose M., Hauf J., Koetter F.,
RA Bernier S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villerioel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay W., Lennard N., McAlay K., Mayes R.,
RA Petit A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
RA Borkova D., Bloeker H., Scharf M., Grimm M., Loeubert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandjean K., Danner D., Herzl A.,
RA Neumann S., Ayrizrau A., Vitale D., Liguori R., Piravandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefor F., Cooke R., Berger C., Montfort A., Caracubeta E.,
RA Gibbons T., Weber N., Vandenbol C., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biske C.,
RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dehla N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sektion W., Murray J., Saeet P., Cordes M., Abu-Threideh J.,

```

```

RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Lactelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Krimer J., Fulton L., Martin B., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Spieh J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali T., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Antonov B., Zidant M., Strong C., Sun H., Lamar B., Jordan C.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez W., Hoffman U., Tili S.,
RA Granat S., Shohy N., Martignac A., Hamed A., Iochi M., Johnson A.,
RA Chen E., Marra M.A., Martenssen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana";
RL Nature 402:769-777(1999);
CC - FUNCTION: Transcription factor. Interacts specifically with the W
box (5'-(T)GAC(C/T)-3'), a frequently occurring elicitor-
responsive cis-acting element. May act also as a disease
resistance protein with a serine/threonine-protein kinase activity
(BY similarity).
CC - SUBCELLULAR LOCATION: Nuclear (Probable).
CC - SIMILARITY: Belongs to the WRKY group I family.
CC - SIMILARITY: Belongs to the disease resistance X-TIR-NB-LRR-X
family.
CC - SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC - SIMILARITY: Contains 1 NB-ARC domain.
CC - SIMILARITY: Contains 1 PAH (paired amphipathic helix) repeat.
CC - SIMILARITY: Contains 1 protein kinase domain.
CC - SIMILARITY: Contains 2 WRKY domains.
CC - DATABASE: NMR-NB-LRRs;
CC NOTE=Functional and comparative genomics of disease resistance gene
homologs;
CC WWW="http://nblrire.ucdavis.edu".
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
CC -----
DR EMBL, AL049638; CAB40943.1; -.
DR EMBL, AL161533; CAB78245.1; -.
DR PIR, T06609; T06609.
DR HSSP; P02876; 9WGA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR000767; Disease_resist.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR002182; NB-ARC.
DR InterPro: IPR003822; PAH.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR000157; TIR.
DR InterPro: IPR003657; WRKY.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF02671; PAH; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF03106; WRKY; 2.
DR PRINTS; PR00364; DISSEA8R81ST.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE NEG.
DR PROSITE; PS50811; WRKY; 2.
KW plant defense; Transferase; Serine/threonine-protein kinase;
transcription regulation; Nuclear protein; ATP-binding; DNA-binding;
Repeat; leucine-rich repeat.

```

```

FT REPEAT 323 369 PAH.
FT DNA_BIND 462 526 WRKY 1.
FT DNA_BIND 635 700 WRKY 2.
FT DOMAIN 800 1087 NB-ARC.
FT REPEAT 1204 1226 LRR 1.
FT REPEAT 1257 1281 LRR 2.
FT REPEAT 1304 1327 LRR 3.
FT REPEAT 1328 1348 LRR 4.
FT REPEAT 1349 1371 LRR 5.
FT REPEAT 1372 1395 LRR 6.
FT REPEAT 1419 1443 LRR 7.
FT DOMAIN 1626 1877 PROTEIN KINASE.
FT NP_BIND 844 851 ATP (POTENTIAL).
FT NP_BIND 1632 1640 ATP (BY SIMILARITY).
FT BINDING 1654 1654 ATP (BY SIMILARITY).
FT ACT_SITE 1758 1758 BY SIMILARITY.
FT DOMAIN 97 307 GLY-RICH.
FT DOMAIN 35 90 POLY-SER.
FT DOMAIN 87 90 POLY-SER.
FT DOMAIN 980 983 POLY-LEU.
FT DOMAIN 1568 1571 POLY-SER.
SQ SEQUENCE 1895 AA; 210320 MW; 1C19D3EE164C9363 CRC64;

```

Query Match 47.2%; Score 41.5; DB 1; Length 1895;
 Best Local Similarity 60.0%; Pred. No. 70;
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 SHOESTYXGSRGRSG 15
 Db 219 SHEDCT-GPARGRSG 232

```

RESULT 11
SF5_RAT STANDARD; PRT; 825 AA.
AC 063003;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE SE5 antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WiStar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obara K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein SE5 in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; D37934; BAA07153.1; -.
DR PIR; JC4163; JC4163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 8681 MW; AF667F82FD555BDF CRC64;

```

Query Match 46.6%; Score 41; DB 1; Length 825;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GXRGSRGRSGS 19
 Db 589 GRGRGRGRSGS 600

```

RESULT 12
PRO_CRESP STANDARD; PRT; 1015 AA.
ID PRO_CRESP
AC Q00586;
DT 15-JUL-1999 (Rel. 38; Created)
DT 15-JUL-1999 (Rel. 38; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Frequency clock protein.
DE PRO.
OS Creopus spinulosus (Chromocrea spinulosa).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=110619;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4896;
RX MEDLINE=97188515; PubMed=9037100;
RA Lewis M.T., Morgan L.W., Feldman J.F.;
RT "Analysis of frequency (frq) clock gene homologs: evidence for a
RT helix-turn-helix transcription factor.";
RL Mol. Gen. Genet. 253:401-414(1997).
CC -1- FUNCTION: Circadian clock component involved in the generation of
CC biological rhythms, in particular in rhythm stability, period
CC length, and temperature compensation. Behaves as a negative
CC element in circadian transcriptional loop (by similarity).
CC -1- SIMILARITY: BELONGS TO THE PRO FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U25850; AAA68072.1; -.
DR PIR; T42013; T42013.
KW Biological rhythms; Transcription regulation; Nuclear protein.
FT DOMAIN 240 245 POLY-SER.
FT DOMAIN 356 368 POLY-GLN.
FT DOMAIN 443 451 POLY-PRO.
FT DOMAIN 584 588 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 892 913 ASB/GLU-RICH (ACIDIC)
SQ SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;

```

Query Match 46.6%; Score 41; DB 1; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 43;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTYXGSRGRSGS 19
 Db 583 NHRKQKTGHTSDSGSSGN 601

```

RESULT 13
COAC_YEAST STANDARD; PRT; 2233 AA.
ID COAC_YEAST
AC Q00955;
DT 01-APR-1993 (Rel. 25; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
DE PAS3 OR ACC1 OR YNR015C OR N3175.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

```

```

RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2015-2022.
RX MEDLINE=92262474; PubMed=1350093;
RA Al-Feel W., Chitara S.S., Wakil S.J.;
RT "Cloning of the yeast P433 gene and primary structure of yeast
  acetyl-CoA carboxylase."
RL Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Pohli T.M.;
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This protein carries three functions: biotin carboxyl
  carrier protein, biotin carboxylase, and carboxyltransferase.
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
  + malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
  = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: Biotin.
CC -1- ENZYME REGULATION: By phosphorylation.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
  step.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation-
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
DR EMBL; M92156; AAA20073.1; -
DR EMBL; Z71631; CAA96294.1; -
DR PIR; S63347; S63347.
DR HSSP; P24182; IBNC.
DR Germonline; 143361; -.
DR GO; GO:0005789; Cytoplasmic reticulum membrane; IDA.
DR GO; GO:0003989; Fatty acid biosynthesis; IMP.
DR GO; GO:0004075; Fatty acid biosynthesis; IMP.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR Pfam; PF02785; Biotin_carb_C_1.
DR Pfam; PF00354; Biotin_lipoyl_1.
DR Pfam; PF01039; Carboxyl_trans_1.
DR Pfam; PF02789; CPase_L_trans_1.
DR Pfam; PF02786; CPase_L_D2_1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KM Fatty acid biosynthesis; Biotin; ligase; Multifunctional enzyme;
  ATP-binding; Phosphorylation.
KW ATP_BIND.
FT NE_BIND 256
FT ACT_SITE 383
FT BINDING 735
FT DOMAIN 1865
FT CONFLICT 1523
FT CONFLICT 1755
FT CONFLICT 1761
SQ SEQUENCE 2233 AA; 250351 MW; 0A335AAD9B1F8308 CRC64;

Query Match          46.6%; Score 41; DB 1; Length 2233;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Db 1200 HOSSNPGAPDRSGSSAS 1217

RESULT 14
ID RO33 MOUSE STANDARD; PRT; 379 AA.
AC OSBG05; OSBHF8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3).
GN HNRP33 OR HNRP33.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC MGI; taxid=10090;
OK NCBI; taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Bayerisalan D.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=FVB/N; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Dietzenko L., Marisina K., Farmer A.A., Rubin C.M., Hong L.,
RA Stopleton M., Soares M.B., Bonaldo M.P., Casavant J.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bock S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences".
  CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
  CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes (By
    similarity).
  CC -1- ALTERNATIVE PRODUCTS:
    CC Event=Alternative splicing; Named isoforms=2;
    CC Name=1;
    CC IsoId=OSBG05-1; Sequence=Displayed;
    CC Name=2;
    CC IsoId=OSBG05-2; Sequence=VSP_007350;
  CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRP, WHICH ARE BASIC AND
    CC GAY-RICH PROTEINS.
  CC -1- SIMILARITY: Contains 2 RNA recognition motif (RNM) domains.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
    CC the European Bioinformatics Institute. There are no restrictions on its
    CC use by non-profit institutions as long as its content is in no way
    CC modified and this statement is not removed. Usage by and for commercial
    CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
    CC or send an email to license@isb-sib.ch).
    CC -----
DR EMBL; AF463524; AAN76922.1; -
DR EMBL; BC023828; AAH23828.1; -
DR EMBL; BC023908; AAH23908.1; -
DR EMBL; BC038364; AAH38364.1; -
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 2.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-6
Sequence: 1 SH0ESTYXGSRGSRGSRGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	797	4	Q16824
2	84	95.5	990	4	Q15206
3	84	95.5	1218	4	Q05331
4	78	88.6	798	4	Q094U3
5	78	88.6	1084	4	Q01212
6	69	78.4	591	4	Q01720
7	69	78.4	687	4	Q094U2
8	67	76.1	322	4	Q05370
9	64	72.7	465	4	Q03838
10	49	55.7	822	3	Q09312
11	47	53.4	96	15	Q08JBU5
12	47	53.4	96	15	Q09WRF4
13	47	53.4	96	15	Q08JBU4
14	47	53.4	96	15	Q08JBU5
15	47	53.4	399	4	Q09NTA9
16	47	53.4	467	4	Q09BWT8

17	47	53.4	700	13	Q042378
18	47	53.4	715	13	Q080F00
19	47	53.4	715	13	Q080G68
20	47	53.4	716	13	Q042107
21	47	53.4	737	4	Q09BQ39
22	47	53.4	843	12	Q090N52
23	46	52.3	96	15	Q08ADX3
24	46	52.3	243	10	Q09AX45
25	46	52.3	426	4	Q08W776
26	45	51.1	669	11	Q08BRU5
27	45	51.1	734	11	Q09SM79
28	44.5	50.6	486	10	Q07XQ91
29	44	50.0	96	15	Q09WSQ0
30	44	50.0	96	15	Q09W9K9
31	44	50.0	96	15	Q090CLO
32	44	50.0	96	15	Q09WSP4
33	44	50.0	96	15	Q0998H5
34	44	50.0	96	15	Q072615
35	44	50.0	96	15	Q090CK1
36	44	50.0	96	15	Q09WRG0
37	44	50.0	96	15	Q09WR27
38	44	50.0	96	15	Q08UMH2
39	44	50.0	96	15	Q091MJ5
40	44	50.0	96	15	Q096H4
41	44	50.0	96	15	Q072610
42	44	50.0	96	15	Q090DY9
43	44	50.0	96	15	Q09GRX5
44	44	50.0	96	15	Q090D29
45	44	50.0	103	12	Q09WDD4

ALIGNMENTS

RESULT 1
Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9106347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene [published erratum appears in Biochemistry 1991 Jun
11;30(23):5814].";
RT Biochemistry 29:9432-9440(1990).
RL EMBL: M60502; AAA63248.1; -
DR GO:00005198; F:structural molecule activity; IEA.
DR InterPro:IPR003303; Flaggrin.
DR Pfam: PF05516; Flaggrin; 4.
DR PRINTS: PR00487; FLAGGRIN.
FT NON_TER 1
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDA86B CRC64;

Query Match 95.5%; Score 84; DB 4; Length 797;
Best local similarity 89.5%; Pred. No. 1.2e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SH0ESTYXGSRGSRGSRGS 19
DB 427 SH0ESTYXGSRGSRGSRGS 445

RESULT 2

```

015206
ID 015206 PRELIMINARY; PRT; 990 AA.
AC 015206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AAA63244.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 990
FT NON_TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A9:991 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGXSRRGSRGSGS 19
Db 551 SHOESTRGSGRSGRSGS 569

RESULT 3
ID 005331 PRELIMINARY; PRT; 1218 AA.
AC 005331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chidsev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993)
RT -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
RT DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
RT TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; M96943; AAA6487.1; -.
DR PIR; A48118; A48118.
DR HSSP; P02593; ICDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR020448; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF000036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR KMW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KMW Polymorphism.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON_TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 1218;
Best Local Similarity 89.5%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGXSRRGSRGSGS 19
Db 773 SHOESTRGSGRSGRSGS 791

RESULT 4
ID 094U3 PRELIMINARY; PRT; 798 AA.
AC 094U3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE DJ4N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBD databases.
DR EMBL; AL36504; CAC13171.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
DR NON_TER 798
FT NON_TER 798
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 88.6%; Score 78; DB 4; Length 798;
Best Local Similarity 84.2%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGXSRRGSRGSGS 19
Db 428 SHOESTRGSGRSGRSGS 446

RESULT 5
ID 001212 PRELIMINARY; PRT; 1084 AA.

```

```
AC Q01212; Q03840;
PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135840; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467
FT PROPEP 468 474
FT CHAIN 475 >591
FT CA_BIND 19 32
FT CA_BIND 62 73
FT NON_TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75B369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.003;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 449 SHOESTXGRSGRSGSGS 467

RESULT 7
ID Q9H4U2; PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DJ4N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.003;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 64 SHOESTXGRSGRSGSGS 82

RESULT 6
ID Q01720; PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PLACENTA;
RA MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Nirusukeiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```
DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135840; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467
FT PROPEP 468 474
FT CHAIN 475 >591
FT CA_BIND 19 32
FT CA_BIND 62 73
FT NON_TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75B369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.003;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 449 SHOESTXGRSGRSGSGS 467

RESULT 7
ID Q9H4U2; PRELIMINARY; PRT; 687 AA.
AC Q9H4U2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE DJ4N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.003;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 64 SHOESTXGRSGRSGSGS 82

RESULT 6
ID Q01720; PRELIMINARY; PRT; 591 AA.
AC Q01720;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PLACENTA;
RA MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Nirusukeiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```

QY      1 SHOESTXGXSRRGSRGSGS 19
      |||||
      449 SHOESTRGRGSRGSRGSGS 467

RESULT 8
ID 075370 PRELIMINARY; PRT; 322 AA.
AC 075370;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Biddermal filaggrin (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9101527; PubMed=9886436;
RA Girbal-Neuhausser E., Durieux J.J., Arnaud M., Dalbon P., Sebba M.,
RA Vincent C., Simon M., Senhu T., Masson-Bessiere C.,
RA Jolivet-Reynaud C., Jolivet M., Serre G.,
RT "The epitopes targeted by the rheumatoid arthritis-associated
RT anti-filaggrin autoantibodies are posttranslationally generated on
RT various sites of (pro)filaggrin by deamination of arginine residues."
RL J. Immunol. 162:585-584(1999).
DR EMBL; AF043380; AAC23559.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 2.
DR PRINTS; PRO0487; FILAGRIN.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230BFF9E0 CRC64;

Query Match
Best Local Similarity 76.1%; Score 67; DB 4; Length 322;
Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 SHOESTXGXSRRGSRGSGS 18
      |||||
      305 SHOESTRGRGSRGSRGSGS 322

RESULT 9
ID 003838 PRELIMINARY; PRT; 465 AA.
AC 003838;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGRIN (PROFILAGRIN) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.,
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene."
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP REVISIONS
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.,
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene."
RL Biochemistry 30:5814-5814(1991).

```

```

CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
CC REPEATS.
CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE
CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN
CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
DR EMBL; M60499; AAA63246.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 3.
DR PRINTS; PRO0487; FILAGRIN.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 465 AA; 50280 MW; C883744C5E1334097 CRC64;

Query Match
Best Local Similarity 72.7%; Score 64; DB 4; Length 465;
Pred. No. 0.016;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 SHOESTXGXSRRGSRGSGS 19
      |||||
      227 SHOESARGRGRGSRGSGS 245

RESULT 10
ID 09312 PRELIMINARY; PRT; 822 AA.
AC 09312;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Related to nucleolar phosphoprotein.
GN B12F1.10.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aligh V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR005120; Smg-4_UZF3.
DR Pfam; PF00076; trm; 1.
DR Pfam; PF03467; Smg4_UZF3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match
Best Local Similarity 55.7%; Score 49; DB 3; Length 822;
Pred. No. 9.8;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      3 QESTXGXSRRGSRGSGS 19
      |||||
      414 RESASGRTRGRGRGCT 430

RESULT 11

```

08JBUS
ID 08JBUS PRELIMINARY; PRT; 96 AA.
AC 08JBUS;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=00KE_XNH1199;
RA Dowling W.E., Kim B., Mason C.J., Masuna K.Monique., Alam U.,
RA Elson L., Bix D.L., Robb M.L., McCutchan F.E., Carr J.K.;
RT "Forty-one near full length HIV-1 sequences from Kenya reveal an
epidemic of subtype A and A-containing recombinants."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457067; AAN03148.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11300 MW; 4A60A702801D3473 CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

RESULT 12
Q9WRP4 PRELIMINARY; PRT; 96 AA.
ID Q9WRP4;
AC Q9WRP4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96CWR1000.4/96CWR1.4;
RX MEDLINE=99329206; PubMed=10400779;
RA Takehisa J., Zekeng L., Ido E., Yamaguchi-Kabata Y., Mboudjeka I.,
RA Harada Y., Miura T., Kapu L., Hayami M.;
RT "Human immunodeficiency virus type 1 intergroup (M/O) recombination in
cancer.";
RL J. Virol. 73:6810-6820(1999).
DR EMBL; AF097693; AAD41705.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11254 MW; F01C751229CFB9A CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 52.9%; Pred. No. 2.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

RESULT 13

08JBV4
ID 08JBV4 PRELIMINARY; PRT; 96 AA.
AC 08JBV4;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=00KE_XNH1144;
RA Dowling W.E., Kim B., Mason C.J., Masuna K.Monique., Alam U.,
RA Elson L., Bix D.L., Robb M.L., McCutchan F.E., Carr J.K.;
RT "Forty-one near full length HIV-1 sequences from Kenya reveal an
epidemic of subtype A and A-containing recombinants."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF457066; AAN03139.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11319 MW; A23861A05083196B CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

RESULT 14
Q9ADPS PRELIMINARY; PRT; 96 AA.
ID Q9ADPS;
AC Q9ADPS;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Vpr protein.
GN VPR.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=98UG57142;
RX MEDLINE=22375625; PubMed=12487816;
RA Harris K.E., Serwadda D., Sewankambo N., Wabwire F., Kim B.,
RA Kigozi G., Kiwanuka N., Phillips J.B., Meehan M., Lutalo T.,
RA Lase J.R., Merling R., Gray R., Mawer M., Bix D.L., Robb M.L.,
RA McCutchan F.E.;
RT "Among 46 Near Full Length HIV Type 1 Genome Sequences from Rakai
District, Uganda, Subtype D and AD Recombinants Predominate."
RL AIDS Res. Hum. Retroviruses 18:1281-1290(2002).
DR EMBL; AF484512; AAN73753.1; -
DR InterPro; IPR000012; Retrov_Vpr/X.
DR Pfam; PF00522; VPR; 1.
DR PRINTS; PR00444; HIVPRVPRX.
KM AIDS.
SQ SEQUENCE 96 AA; 11356 MW; 5EAB1F03A30F5221 CRC64;

Query Match 53.4%; Score 47; DB 15; Length 96;
Best Local Similarity 58.8%; Pred. No. 2.2;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QESTXGXSRRSGRSGS 19
Db 77 QHSRIGITRRGRNGS 93

```

RESULT 15
Q9NTA9
ID Q9NTA9 PRELIMINARY; PRT; 399 AA.
AC Q9NTA9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKF2P761B0333.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Bioecker H., Bioecker M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.,
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137423; CAB7073.1; -.
DR PIR; T46259; T46259.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:Nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00490; HEPICc; 1.
KM Hypothetical protein; ATP-binding; Helicase; Hydrolase.
FT NON TER
SQ SEQUENCE 399 AA; 44075 MW; BC09B3A0E4CE5E96 CRC64;

Query Match 53.4%; Score 47; DB 4; Length 399;
Best Local Similarity 55.6%; Pred. No. 9.9;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Cy 1 SHQESTXGXRGRSGRSG 18
Db 336 SNGRSGWSSGRSGRSG 353

```

Search completed: September 28, 2004, 06:12:45
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-6
Perfect score: 88
Sequence: 1 SHQESTXGXSRRGSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	AAW61515	Peptide c
2	84	95.5	19	AAW61506	Peptide c
3	84	95.5	19	AAW61517	Peptide c
4	84	95.5	19	AAW61510	Peptide c
5	84	95.5	19	AAW61505	Peptide c
6	84	95.5	19	AAW61514	Peptide c
7	84	95.5	19	AAW61516	Peptide c
8	84	95.5	1467	ABB97605	Novel hum
9	78	88.6	19	AAW61508	Peptide c
10	78	88.6	19	AAW61507	Peptide c
11	78	88.6	19	AAW61512	Peptide c
12	78	88.6	19	AAW61509	Peptide c
13	78	88.6	19	AAW61511	Peptide c
14	78	88.6	19	AAW61513	Peptide c
15	75	85.2	21	AAW61520	Peptide c
16	75	85.2	22	AAW61520	Peptide c
17	53	60.2	330	AAW61520	Peptide c
18	53	60.2	330	AAW61520	Peptide c
19	53	60.2	330	AAW61520	Peptide c
20	49	55.7	330	AAW61520	Peptide c
21	47	53.4	443	AAW61520	Peptide c
22	47	53.4	574	AAW61520	Peptide c
23	47	53.4	700	AAW61520	Peptide c
24	47	53.4	752	AAW61520	Peptide c
25	45	51.1	1711	AAW61520	Peptide c

26	45	51.1	1951	4	AAW78835	Human pro
27	45	51.1	2348	4	ABG10929	Novel hum
28	44	50.0	477	6	ABO07142	Novel hum
29	44	50.0	843	3	AAV54044	Amino aci
30	44	50.0	1199	4	ABBS8274	Drosophi
31	43	48.9	369	4	ABG04454	Novel hum
32	43	48.9	370	4	ABG05066	Novel hum
33	43	48.9	477	4	ABG08221	Novel hum
34	43	48.9	641	4	ABG19110	Novel hum
35	43	48.9	694	2	AAW1267	Drosophi
36	43	48.9	694	2	AAW1267	Drosophi
37	43	48.9	797	5	ABP73371	Candida
38	42.5	48.3	261	4	AAU65828	Propionib
39	42.5	48.3	261	6	ABM62347	Propionib
40	42.5	48.3	353	6	ABM65924	Propionib
41	42	47.7	34	4	AAW17461	Peptide #
42	42	47.7	34	4	ABB16481	Peptide #
43	42	47.7	34	4	AAW29981	Peptide #
44	42	47.7	34	4	ABB31273	Peptide #
45	42	47.7	34	4	ABB21823	Protein #

ALIGNMENTS

RESULT 1
ID AAW61515 standard; peptide; 19 AA.
AC AAW61515;
DT 26-OCT-1998 (first entry)
DE Peptide cFA, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
XX MO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPJ; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX Sequence 19 AA;

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 6e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGXSRRGRSGSGS 19
 |||||
 1 SHOESTAGRSRRGRSGSGS 19

RESULT 2

AAW61506 standard; peptide; 19 AA.

AAW61506;

26-OCT-1998 (first entry)

Peptide cfc2, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 solid phase synthesis; peptide amide; polyclonal antibody;
 monoclonal antibody.

Synthetic.
 Homo sapiens.

Key Location/Qualifiers

Modified-site 9 /note="Citrulline"

W09822503-A2.

28-MAY-1998.

14-NOV-1997; 97WO-NI000624.

15-NOV-1996; 96NL-01004539.

(SCHE-) STICHTING SCHEIKUNDDIG ONDERZOEK IN NEDER.

(TEWE-) STICHTING TECH WETENSCHAPPEN.

Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

WPI; 1998-398613/34.

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.
 The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the
 detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 8.7e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGXSRRGRSGSGS 19
 |||||
 1 SHOESTRGSRRGRSGSGS 19

RESULT 3

AAW61517 standard; peptide; 19 AA.

AAW61517;

26-OCT-1998 (first entry)

Peptide cfcQ, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 solid phase synthesis; peptide amide; polyclonal antibody;
 monoclonal antibody.

Synthetic.
 Homo sapiens.

W09822503-A2.

28-MAY-1998.

14-NOV-1997; 97WO-NI000624.

15-NOV-1996; 96NL-01004539.

(SCHE-) STICHTING SCHEIKUNDDIG ONDERZOEK IN NEDER.

(TEWE-) STICHTING TECH WETENSCHAPPEN.

Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

WPI; 1998-398613/34.

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.
 The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the
 detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 8.7e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGXSRRGRSGSGS 19
 |||||
 1 SHOESTGSRGRSGSGS 19

RESULT 4

AAW61510 standard; peptide; 19 AA.

AAW61510;

26-OCT-1998 (first entry)

Peptide cfc6, based on cDNA of a profilaggrin repeat.

Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 solid phase synthesis; peptide amide; polyclonal antibody;
 monoclonal antibody.

Synthetic.
 Homo sapiens.

[illegible]

```

PF 14-NOV-1997; 97WO-NL000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;
DR WPI; 1998-398613/34.
XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure, Page 6; 19pp; English.
XX
CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the proflilagrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
XX
SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.7e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQESTXGXSRSRGRSGSGS 19
   |||||
Db 1 SHQESTXGSRGRSGRSGSGS 19

RESULT 6
AAW61514
ID AAW61514 standard; peptide; 19 AA.
XX
AC AAW61514;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cf, based on cDNA of a proflilagrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KM monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9822503-A2.
XX
PD 28-MAY-1998.
XX
PF 14-NOV-1997; 97WO-NL000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
PI Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;
DR WPI, 1998-398613/34.
XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX

```

PS Disclosure; Page 6; 19pp; English.

XX CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.7e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYXGSRGRSGSGS 19

DB 1 SHQESTYXGSRGRSGSGS 19

RESULT 7

AAW61516

ID AAW61516 standard; peptide; 19 AA.

XX AC AAW61516;

XX DT 26-OCT-1998 (first entry)

XX DE Peptide cFE, based on cDNA of a proflilagrin repeat.

XX KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX KM solid phase synthesis; peptide amide; polyclonal antibody;

XX KM monoclonal antibody.

XX OS Synthetic.

OS Homo sapiens.

XX PN W09822503-A2.

XX PD 28-MAY-1998.

XX PF 14-NOV-1997; 97MO-NL000624.

XX PR 15-NOV-1996; 96NL-01004539.

XX PA (SCHE-) STICHTING SCHEIKONDIC ONDERZOEK IN NEDER.

XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX PI Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;

XX DR WPI; 1998-398613/34.

XX PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX PS Disclosure; Page 6; 19pp; English.

XX SQ Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.7e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYXGSRGRSGSGS 19

DB 1 SHQESTYXGSRGRSGSGS 19

RESULT 8

ABB97605

ID ABB97605 standard; protein; 1467 AA.

XX AC ABB97605;

XX DT 27-JUN-2002 (first entry)

XX DE Novel human protein SEQ ID NO: 873.

XX KM Human; anti-neoplastic; vulnary; anti-inflammatory; immunomodulator;

XX KM antineoplastic; cerebroprotective; cytoskeletal; rheumatic; gene therapy;

XX KM neuroprotective; antiparkinsonian; protein therapy; EST;

XX KM expressed sequence tag.

XX OS Homo sapiens.

XX PN W0200222660-A2.

XX PD 21-MAR-2002.

XX PF 10-SEP-2001; 2001MO-US026015.

XX PR 11-SEP-2000; 2000US-00659671.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX PI Xue AJ, Yang Y, Wehrman T, Dymnac RT;

XX DR WPI; 2002-292408/33.

XX PF N-PSDB; AEN32791.

XX PT An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

XX PS Example 2; SEQ ID NO 873; 509pp; English.

XX CC The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

XX SQ Sequence 1467 AA;

Query Match 95.5%; Score 84; DB 5; Length 1467;

Best Local Similarity 89.5%; Pred. No. 6e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYXGSRGRSGSGS 19

DB 773 SHQESTYXGSRGRSGSGS 791

RESULT 9

AAW61508

ID AAW61508 standard; peptide; 19 AA.

XX AC AAW61508;

DT 26-OCT-1998 (first entry)
 XX Peptide cfc4, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 8.2e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHOESTXGXRGRSGSGS 19
 DB 1 SHOESTRGRXRGRSGSGS 19
 RESULT 10
 AAM61507
 ID AAM61507 standard; peptide; 19 AA.
 AC AAM61507;
 XX
 XX 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc3, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 11
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX
 XX Disclosure; Page 6; 19pp; English.
 XX
 XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 8.2e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHOESTXGXRGRSGSGS 19
 DB 1 SHOESTRGRXRGRSGSGS 19
 RESULT 11
 AAM61512
 ID AAM61512 standard; peptide; 19 AA.
 AC AAM61512;
 XX
 XX 26-OCT-1998 (first entry)
 XX
 XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7
 FT /note= "Citruiline"
 FT Modified-site 13
 FT /note= "Citruiline"
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX

```

XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX SO Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGSRGSRGSGS 19
DB 1 SHOESTXGSRGSRGSGS 19

RESULT 12
AAM61509
ID AAM61509 standard; peptide; 19 AA.
XX AAM61509;
XX 26-OCT-1998 (first entry)
XX Peptide cfc5, based on cDNA of a profilaggrin repeat.
XX DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclonal antibody.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 16 /note= "Citruilline"
XX FT
XX PN WO9822503-A2.
XX 28-MAY-1998.
XX PD
XX PI 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX SO Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
Best Local Similarity 84.2%; Pred. No. 8.2e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGSRGSRGSGS 19
DB 1 SHOESTXGSRGSRGSGS 19

RESULT 13
AAM61511
ID AAM61511 standard; peptide; 19 AA.
XX AAM61511;
XX 26-OCT-1998 (first entry)
XX DE Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KM solid phase synthesis; peptide amide; polyclonal antibody;
XX KM monoclonal antibody.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 7 /note= "Citruilline"
XX FT Modified-site 11 /note= "Citruilline"
XX FT
XX PN WO9822503-A2.
XX 28-MAY-1998.
XX PD
XX PI 14-NOV-1997; 97WO-NL000624.
XX PR 15-NOV-1996; 96NL-01004539.
XX PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX PI Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA;
XX DR WPI; 1998-398613/34.
XX PT Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX PS Disclosure; Page 6; 19pp; English.
XX CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

```

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGXRGRSGSGS 19
 |||||
 DB 1 SHOESTXGXRGRSGSGS 19

RESULT 14

AAW61513
 ID AAW61513 standard; peptide; 19 AA.

AC AAW61513;

DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citruilline"

FT Modified-site 16 /note= "Citruilline"

FT Modified-site 16 /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij MWM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 8.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGXRGRSGSGS 19
 |||||
 DB 1 SHOESTXGXRGRSGSGS 19

RESULT 15

AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;

DT 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note= "Citruilline"

FT Modified-site 9 /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij MWM, Schellekens GA, Raats JMH, Hoet RMA;

XX WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;

Best Local Similarity 88.9%; Pred. No. 2.4e-05;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HOESTXGXRGRSGSGS 19
 |||||
 DB 4 HOESTXGXRGRSGSGS 21

Search completed: September 28, 2004, 06:24:16
 Job time : 55.1302 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-6

Sequence: 1 SHQESTXGKSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
 - 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/1/pubppaa/US10C_NEW_PUB.pep:*
 - 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
 - 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-2
4	84	95.5	19	9	US-09-308-150-6
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	78	88.6	19	9	US-09-308-150-3
9	78	88.6	19	9	US-09-308-150-4
10	78	88.6	19	9	US-09-308-150-5
11	78	88.6	19	9	US-09-308-150-7
12	78	88.6	19	9	US-09-308-150-8
13	78	88.6	19	9	US-09-308-150-9
14	75	85.2	21	9	US-09-308-150-10
15	75	85.2	22	9	US-09-747-029A-22

16	49	55.7	506	16	US-10-437-963-185974	Sequence 185974,
17	46	52.3	243	16	US-10-437-963-103033	Sequence 103033,
18	44.5	50.6	181	16	US-10-437-963-117640	Sequence 117640,
19	44	50.0	155	12	US-10-425-114-40476	Sequence 40476, A
20	44	50.0	345	12	US-10-425-114-62866	Sequence 62866, A
21	44	50.0	430	16	US-10-437-963-164039	Sequence 164039,
22	44	50.0	477	15	US-10-161-927-62	Sequence 62, Appl
23	44	50.0	653	9	US-09-746-801A-13	Sequence 13, Appl
24	44	50.0	730	14	US-10-359-431-19	Sequence 49, Appl
25	44	50.0	842	14	US-10-359-431-50	Sequence 50, Appl
26	44	50.0	842	14	US-10-359-431-51	Sequence 51, Appl
27	44	50.0	843	14	US-10-209-264-7	Sequence 2, Appl
28	44	50.0	843	14	US-10-359-431-45	Sequence 45, Appl
29	44	50.0	843	14	US-10-359-431-48	Sequence 48, Appl
30	44	50.0	843	14	US-10-359-431-59	Sequence 59, Appl
31	44	50.0	845	14	US-10-359-431-16	Sequence 46, Appl
32	44	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
33	44	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
34	44	50.0	1087	16	US-10-641-991-74	Sequence 24, Appl
35	43	48.9	232	16	US-10-437-963-183394	Sequence 183394,
36	43	48.9	247	16	US-10-424-599-187446	Sequence 187446,
37	43	48.9	455	16	US-10-437-963-185897	Sequence 185897,
38	43	48.9	484	16	US-10-437-963-185432	Sequence 185432,
39	43	48.9	488	16	US-10-437-963-146097	Sequence 146097,
40	43	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
41	43	48.9	694	14	US-10-157-548-2	Sequence 2, Appl
42	43	48.9	797	14	US-10-032-585-7208	Sequence 7208, Ap
43	42.5	48.3	92	16	US-10-437-963-118183	Sequence 118183,
44	42	47.7	34	9	US-09-864-761-37121	Sequence 37121, A
45	42	47.7	179	12	US-10-425-114-53545	Sequence 53545, A

ALIGNMENTS

US-09-308-150-12
; Sequence 12, Application US/09308.150
; Patent No. US20020137092A1
GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Kaats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-12
Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 85.5%; Pred. No. 5.3e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHQESTXGKSRGRSGSGS 19

Db 1 SH0ESTXGSRGSRGSRGS 19

RESULT 2

US-09-308-150-1
Sequence 1, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 7.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTXGSRGSRGSRGS 19
Db 1 SH0ESTXGSRGSRGSRGS 19

RESULT 3

US-09-308-150-2
Sequence 2, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 7.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTXGSRGSRGSRGS 19
Db 1 SH0ESTXGSRGSRGSRGS 19

RESULT 4

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SH0ESTXGSRGSRGSRGS 19
Db 1 SH0ESTXGSRGSRGSRGS 19

RESULT 5

US-09-308-150-11
Sequence 11, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 11
LENGTH: 19

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-11

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1 SHOEETXGXSRRGRSGSGS 19
Db      1 SHOEETGRSRGRSGSGS 19

```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-13

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1 SHOEETXGXSRRGRSGSGS 19
Db      1 SHOEETGRSRGRSGSGS 19

```

```

RESULT 7
US-09-308-150-14

```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-14

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 7.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1 SHOEETXGXSRRGRSGSGS 15
Db      1 SHOEETGRSRGRSGSGS 19

```

```

RESULT 8
US-09-308-150-3
; Sequence 3, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; US-09-308-150-3

```

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.9e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSRGSGS 19
Db 1 SH0ESTRGRSRGSRGSGS 19

RESULT 9

US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.9e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSRGSGS 19
Db 1 SH0ESTRGRSRGSRGSGS 19

RESULT 10

US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.9e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSRGSGS 19
Db 1 SH0ESTRGRSRGSRGSGS 19

RESULT 11

US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SH0ESTXGXSRRGSRGSGS 19
Db 1 SH0ESTRGRSRGSRGSGS 19

RESULT 12

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-8

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308.150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Walthervus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-9

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.9e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGSRGSRGSGS 19
Db 1 SHQSTXGSRGSRGSGS 19

RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308.150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Walthervus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
NAME/KEY: DISULFID
LOCATION: (3) .. (16)
US-09-308-150-10

Query Match 85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.00018;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGSRGSRGSGS 19
Db 4 HOESTXGSRGSRGSGS 21

RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydia
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.0031NPOS00 INNS:031
CURRENT APPLICATION NUMBER: US/09/747.029A
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00670195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptides
NAME/KEY: MOD_RES
LOCATION: (10)..(10)
OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

Query Match 85.2%; Score 75; DB 9; Length 22;
Best Local Similarity 88.9%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGXSRRGRCGRSGS 19
|||
Db 5 HOESTXGXSRRGRCGRSGS 22

Search completed: September 28, 2004, 07:28:51
Job time : 112.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-6
Perfect score: 8
Sequence: 1 SHQSTXKXSRGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgcn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgcn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgcn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgcn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgcn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgcn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	54.5	421	4	US-09-252-991A-22326
2	44	50.0	483	4	US-09-252-991A-19015
3	44	50.0	653	4	US-09-513-057C-13
4	44	50.0	730	4	US-08-591-502B-49
5	44	50.0	842	4	US-08-591-502B-50
6	44	50.0	842	4	US-08-591-502B-51
7	44	50.0	843	4	US-09-719-528A-2
8	44	50.0	843	4	US-08-591-502B-45
9	44	50.0	843	4	US-08-591-502B-48
10	44	50.0	843	4	US-08-591-502B-59
11	44	50.0	845	4	US-08-591-502B-46
12	43	48.9	169	3	US-09-342-084-6
13	42	47.7	138	4	US-09-252-991A-16961
14	42	47.7	786	4	US-09-252-991A-24514
15	42	47.7	722	4	US-09-984-890-4
16	42	47.7	724	4	US-09-984-890-2
17	42	47.7	745	4	US-09-523-849-36
18	41	46.6	255	4	US-09-252-991A-19598
19	41	46.6	258	4	US-09-252-991A-18812
20	41	46.6	406	4	US-09-252-991A-19857
21	41	46.6	409	4	US-09-489-039A-14217
22	41	46.6	464	4	US-09-252-991A-31219
23	41	46.6	722	4	US-08-817-832B-32
24	41	46.6	2237	1	US-08-354-973-1
25	40.5	46.0	258	4	US-10-164-593-59
26	40.5	46.0	274	4	US-09-976-594-417
27	40.5	46.0	378	4	US-10-164-593-2

28	40	45.5	120	4	US-09-252-991A-17025	Sequence 17025, A
29	40	45.5	197	4	US-09-252-991A-23579	Sequence 23579, A
30	40	45.5	274	4	US-09-252-991A-16696	Sequence 16696, A
31	40	45.5	465	4	US-09-252-991A-18919	Sequence 18919, A
32	40	45.5	486	2	US-08-821-355A-8	Sequence 8, Appl1
33	40	45.5	486	2	US-09-003-687A-8	Sequence 8, Appl1
34	40	45.5	486	2	US-09-136-605-8	Sequence 8, Appl1
35	40	45.5	508	2	US-08-818-024-3	Sequence 3, Appl1
36	40	45.5	508	3	US-09-334-775A-3	Sequence 3, Appl1
37	40	45.5	660	3	US-09-058-489-18	Sequence 18, Appl1
38	40	45.5	660	3	US-09-058-489-91	Sequence 91, Appl
39	40	45.5	660	4	US-09-976-594-787	Sequence 787, Appl
40	40	45.5	708	4	US-09-252-991A-16074	Sequence 18074, A
41	39.5	44.9	160	4	US-09-252-991A-26901	Sequence 26901, A
42	39	44.3	176	4	US-09-252-991A-29512	Sequence 29512, A
43	39	44.3	195	4	US-09-252-991A-30082	Sequence 30082, A
44	39	44.3	204	4	US-09-252-991A-21317	Sequence 21317, A
45	39	44.3	468	4	US-09-252-991A-24394	Sequence 24394, A

ALIGNMENTS

```
RESULT 1
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326
Query Match      54.5%; Score 48; DB 4; Length 421;
Best Local Similarity 52.9%; Pred. No. 4.6;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Cy      2 HOESTXKXSRGRSG 18
Db      13 HARTGTGRRGRSG 29
RESULT 2
US-09-252-991A-19015
; Sequence 19015, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19015
; LENGTH: 483
; TYPE: PRT
```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19015

Query Match 50.0%; Score 44; DB 4; Length 483;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 H0ESTXGSRGSRGSGS 18
DB 94 HAQCDPGAGTGRAGRAG 110

RESULT 3
US-09-513-057C-13
Sequence 13, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 13
LENGTH: 653
TYPE: PRT
ORGANISM: Cardamine oligosperma
US-09-513-057C-13

Query Match 50.0%; Score 44; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 SH0ESTXGSRGSRGSGS 19
DB 573 SRQVSTASASGREGISGS 591

RESULT 4
US-08-591-502B-49
Sequence 49, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08665
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-591-502B-49

Query Match 50.0%; Score 44; DB 4; Length 730;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSJARGRGSRGSGS 242

RESULT 5
US-08-591-502B-50
Sequence 50, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08665
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-08-591-502B-50

Query Match 50.0%; Score 44; DB 4; Length 842;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGSGS 242

RESULT 6
US-08-591-502B-51
Sequence 51, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 842 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-08-591-502B-51

US-08-591-502B-51

Query Match 50.0%; Score 44; DB 4; Length 842;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGSGS 242

RESULT 7
US-09-719-528A-2
Sequence 2, Application US/09719528A
Patent No. 6558675
GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi
Chen, Wei Ning
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladak & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10023

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/719,528A
FILING DATE: 30-Apr-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SG98/00046
FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:
NAME: Mass, Clifford J.
REGISTRATION NUMBER: 30,086
REFERENCE/DOCKET NUMBER: U-013109-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 708-1800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2

Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARGSGSGS 242

RESULT 8
US-08-591-502B-45
Sequence 45, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:

APPLICANT: Chisari, Francis V
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45
Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QESTXKSRGRSGSGS 19
Db 226 QPQGSIMARKSGSGS 242
RESULT 9
US-08-591-502B-48
Sequence 48, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870

FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-08-591-502B-48
Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 QESTXKSRGRSGSGS 19
Db 226 QPQGSIMARKSGSGS 242
RESULT 10
US-08-591-502B-59
Sequence 59, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Query Match 50.0%; Score 44; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGSRGSGS 19
DB 226 QPQGSILARKSGRSGS 242

RESULT 11
US-08-591-502B-46
Sequence 46, Application US/08591502B
Patent No. 6607727
GENERAL INFORMATION:
APPLICANT: Chisari, Francis V.
TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
Lymphocyte Responses to Hepatitis B Virus
NUMBER OF SEQUENCES: 99
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08665
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 845 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16..18
OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = unknown"
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-08-591-502B-46

Query Match 50.0%; Score 44; DB 4; Length 845;

Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 3 QESTXGSRGSRGSGS 19
DB 228 QPQGSILARKSGRSGS 244

RESULT 12
US-09-342-084-6
Sequence 6, Application US/09342084
Patent No. 6251668
GENERAL INFORMATION:
APPLICANT: Canoon, Rebecca E.
APPLICANT: Sakai, Hajime
TITLE OF INVENTION: Transcription Coactivators
FILE REFERENCE: BB-1169-A
CURRENT APPLICATION NUMBER: US/09/342,084
CURRENT FILING DATE: 1999-06-29
EARLIER APPLICATION NUMBER: 60/092,659
EARLIER FILING DATE: July 13, 1998
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: UNSURE
LOCATION: (128)
US-09-342-084-6

Query Match 48.9%; Score 43; DB 3; Length 169;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTXGSRGSRGSGS 19
DB 15 NNKXSGSGSRGSRGSGS 33

RESULT 13
US-09-252-991A-16961
Sequence 16961, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16961
LENGTH: 138
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16961

Query Match 47.7%; Score 42; DB 4; Length 138;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 8 GXSRGSRGSGS 19
DB 106 GPGRGGRGGRGSGS 117

RESULT 14

```

US-09-252-991A-24514
; Sequence 24514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24514
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24514

```

```

Query Match          47.7%; Score 42; DB 4; Length 586;
Best Local Similarity 72.7%; Pred. No. 56;
Matches      8; Conservative      0; Mismatches      3; Indels      0; Gaps      0;

```

```

QY      8 GXSGRSGRSG 18
      |||||
DB      252 GGGGRSGRSG 262

```

```

RESULT 15
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4

```

```

Query Match          47.7%; Score 42; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 69;
Matches      9; Conservative      1; Mismatches      5; Indels      0; Gaps      0;

```

```

QY      5 STYXGRSGRSGS 19
      |||||
DB      562 SPGHSGRSGS 576

```

```

Search completed: September 28, 2004, 06:26:44
Job time : 15.4479 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-7

Perfect score: 88

Sequence: 1 SHOESTXGRSXRGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	416	2 A32947	filaggrin precursor
2	85	96.6	2248	2 A35938	profilaggrin - hum
3	76	86.4	591	2 A45135	profilaggrin - hum
4	52	59.1	506	1 W2ML47	E2 protein - human
5	49	55.7	822	2 T51049	related to nucleol
6	48	54.5	471	2 T33997	hypothetical prote
7	47	53.4	306	2 T21220	hypothetical prote
8	47	53.4	399	2 T16259	hypothetical prote
9	44	50.0	123	2 T16234	hypothetical prote
10	44	50.0	800	2 T02852	probable membrane
11	43.5	49.4	286	2 S50855	neurotrophin-6 - s
12	43	48.9	649	2 G71283	probable ATP-depen
13	43	48.9	694	2 S71782	wingless receptor
14	43	48.9	849	2 A96592	hypothetical prote
15	43	48.9	1015	2 B42013	frequency clock pr
16	43	48.9	1507	2 B47328	natural killer cel
17	42	47.7	136	2 T35632	hypothetical prote
18	42	47.7	1119	2 T16720	hypothetical prote
19	42	47.7	1337	2 T30291	dehydrogenase - strep
20	42	47.7	2420	2 A84652	hypothetical prote
21	41	46.6	625	2 A34615	profilaggrin - rat
22	41	46.6	745	2 G01025	serine/threonine p
23	41	46.6	774	2 I48609	probable serine/th
24	41	46.6	825	2 JC4163	DNA-binding protei
25	41	46.6	836	2 G84727	probable DNA topoi
26	41	46.6	2233	2 S63347	acetyl-CoA carboxy
27	40	45.5	167	2 S05545	dehydrin 3 - maize
28	40	45.5	167	2 T52599	squamosa promoter
29	40	45.5	168	2 A39316	RAB-17 protein - m

30	40	45.5	168	2 S08633	RAB-17 protein - m
31	40	45.5	174	2 T52600	squamosa promoter
32	40	45.5	195	2 T37163	hypothetical prote
33	40	45.5	197	2 T18818	hypothetical prote
34	40	45.5	203	2 C87801	protein C10G11.9
35	40	45.5	312	2 A31846	130K paracrystall
36	40	45.5	315	2 S53589	SOL2 protein - yea
37	40	45.5	373	2 T02976	probable DNA bindi
38	40	45.5	454	2 A36643	protein kinase (EC
39	40	45.5	482	2 C86322	hypothetical prote
40	40	45.5	484	2 S53641	protein kinase clk
41	40	45.5	508	2 T22440	hypothetical prote
42	40	45.5	526	2 A41698	cell division cont
43	40	45.5	638	2 I53169	cytokestatin 2 - hu
44	40	45.5	692	2 H71426	hypothetical prote
45	40	45.5	760	2 F75530	ribonuclease - Del

ALIGNMENTS

RESULT 1
A32947
filaggrin precursor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999
C/Accession: A32947
R/McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A/Title: Characterization of a cDNA clone encoding human filaggrin and localization of th
A/Reference number: A32947; MUID:89286901; PMID:2740331
A/Accession: A32947
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-416 <MCK>
A/Cross-references: GB:W24355; NID:9182604; PIDN:AA52454.1; PID:9182605
A/Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C/Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 96.6%; Score 85; DB 2; Length 416;
Best Local Similarity 89.5%; Pred. No. 4.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

cy 1 SHOESTXGRSXRGRSGS 19
db 7 SHOESTXGRSXRGRSGS 25

RESULT 2
A35938
profilaggrin - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C/Accession: A35938
R/Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A/Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A/Reference number: A35938; MUID:91064347; PMID:2248957
A/Accession: A35938
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-2248 <GAN>
A/Cross-references: GB:U02929
C/Genetics:
A/Gene: GDB:FLG
A/Cross-references: GDB:119912; OMIM:135940
A/Map position: 1q21-1q21
C/Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology


```

RESULT 10
T02852
Probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: T02852; H81462
R:Wiley, P.J.
Submitted to the EMBL Data Library, May 1998
A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.
A:Reference number: Z14740
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: T02852
A:Molecule type: DNA
A:Residues: 1-800 <MYL>
A:Cross-references: EMBL:AE001274; NID:g3264850; PID:g2266920
R:Wiley, P.J.; Ardaman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F
Proc. Natl. Acad. Sci. U.S.A. 96 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-co
A:Reference number: A81455; WUID:99178997; PMID:10077609
A:Accession: H81462
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-800 <PYL>
A:Cross-references: GB:AE001274; NID:g3264850; PIDN:AA61675.1; PID:g2266920; GSPDB:GN000
C:Experimental source: strain WHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L1439.4
A:Map position: 1

Query Match          50.0%; Score 44; DB 2; Length 800;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      2 HOESTXGRSGXGRSGSGS 19
          ::|||::|||
Db      429 HRDVGRLSTTAGRSGS 446

RESULT 11
S50855
Neurotrophin-6 - southern platyfish
C:Species: Xiphophorus maculatus (southern platyfish)
C:Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #ext_change 20-Jan-2003
C:Accession: S50855
R:Geetz, R.; Koester, R.; Winkler, C.; Raulf, F.; Lottspeich, F.; Scharlt, M.; Thoenen, I
Nature 372, 266-269, 1994
A:Title: Neurotrophin-6 is a new member of the nerve growth factor family.
A:Reference number: S50855; WUID:95059452; PMID:796471
A:Accession: S50855
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <GOE>
A:Cross-references: GB:U36942; NID:g642673; PIDN:AAA61923.1; PID:g642674
C:Superfamily: nerve growth factor beta chain

Query Match          49.4%; Score 43.5; DB 2; Length 286;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      1 SHOES---TXGRSGRSGRSGS 19
          ::|||::|||
Db      203 THRSGVIGRSGRSGRSGSGS 224

RESULT 12
G71283
Probable ATP-dependent RNA helicase - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #ext_change 02-Feb-2001
C:Accession: G71283
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

```

erson, J.; Khajak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDev
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A/Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A/Reference number: A71250; MIMD:98332770; PMID:9665876
A/Accession: G71283
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-649 <COL>
A/Cross-references: GB:AE001248; GB:AE000520; NID:g3232074; PIDN:AA065738.1; PID:g323207
A/Experimental source: strain Nichols
C/Genetics:
A/Gene: TP0770
C/Keywords: ATP; nucleotide binding; P-loop
F:91-98/Region: nucleotide-binding motif A (P-loop)
F:192-197/Region: nucleotide-binding motif B
F:196-199/Region: DEAD motif

Query Match 48.9%; Score 43; DB 2; Length 649;
Best Local Similarity 44.4%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQSTXGRSGRSGS 19
DB 364 HDSATYTHRYGRTGRAGS 381

RESULT 13

Wingless receptor precursor dfz2 - fruit fly (*Drosophila melanogaster*)
C/Species: *Drosophila melanogaster*
C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
C/Accession: S71786; S78444
R/Bhanot, P.; Brink, M.; Harryman Samos, C.; Heish, J.C.; Wang, Y.; Macke, J.P.; Andrew,
Nature 382, 225-230, 1996
A/Title: A new member of the frizzled family from *Drosophila* functions as a wingless rec
A/Reference number: S71786; MIMD:96353971; PMID:8717036
A/Accession: S71786
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-694 <BHA>
A/Cross-references: EMBL:U65589
A/Note: mRNA was also sequenced
R/Bhanot, P.; Wang, Y.; Nathans, J.
submitted to the EMBL Data Library, July 1996
A/Reference number: S78444
A/Accession: S78444
A/Molecule type: DNA
A/Residues: 1-416; T, 418-694 <BHW>
A/Cross-references: EMBL:U65589; NID:g1518050; PIDN:AA047273.1; PID:g1518051
C/Genetics:
A/Gene: dfz2
A/Cross-references: FlyBase:FBgn0016797
C/Superfamily: fruit fly frizzled protein
C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
Best Local Similarity 52.6%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 19
DB 182 SYTEAGSGSGSGSGSGS 200

RESULT 14

A96592
hypothetical protein F14C21.55 [imported] - *Arabidopsis thaliana*
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: A96592
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A/Reference number: A86141; MIMD:21016719; PMID:11130712
A/Accession: A96592
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-849 <STO>
A/Cross-references: GB:AE005173; NID:g11055759; PIDN:AA028231.1; GSPDB:GN00141
C/Genetics:
A/Gene: F14C21.55
A/Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGS 19
DB 736 SSANRSRGGRIGS 750

RESULT 15

frequency clock protein - *Creopus spinulosus*
C/Species: *Creopus spinulosus*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C/Accession: T42013
R/Lewis, M.T.; Morgan, L.W.; Feldman, J.F.
Mol. Gen. Genet. 253, 401-414, 1997
A/Title: Analysis of frequency (frc) clock protein homologs: evidence for a helix-turn-h
A/Reference number: Z22024; MIMD:9718515; PMID:9037100
A/Accession: T42013
A/Status: preliminary; translated from GB/EMBL/DBU
A/Molecule type: DNA
A/Residues: 1-1015 <LEW>
A/Cross-references: EMBL:U25850; NID:g852501; PID:g852502; PIDN:AAA68072.1
C/Genetics:
A/Gene: frc
A/Introns: 100/1

Query Match 48.9%; Score 43; DB 2; Length 1015;
Best Local Similarity 42.1%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGS 19
DB 583 NHRKQKTGSHSGSGSGN 601

Search completed: September 28, 2004, 06:15:01
Job time : 13.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.352 Million cell updates/sec

Title: US-09-308-150-7

Perfect score: 88

Sequence: 1 SHQSTYGRSXGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	416	1	FILA_HUMAN
2	52	59.1	506	1	VE2_HPV47
3	44	50.0	123	1	LSM4_CAEEL
4	43	48.9	694	1	FRZ2_DROME
5	43	48.9	1015	1	FRQ_TREPP
6	43	48.9	1453	1	NKCR_MOUSE
7	42	47.7	1337	1	DEXT_STRDO
8	41	46.6	774	1	MRR2_MOUSE
9	41	46.6	825	1	SE5_RAT
10	41	46.6	2233	1	COAC_YEAST
11	40	45.5	168	1	DH1_MAIZE
12	40	45.5	315	1	SOL2_YEAST
13	40	45.5	484	1	CLK1_HUMAN
14	40	45.5	526	1	2ASA_YEAST
15	40	45.5	638	1	K220_HUMAN
16	40	45.5	1527	1	ARHB_RAT
17	39.5	44.9	135	1	RBP1_DROME
18	39	44.3	102	1	SAAS_MESAU
19	39	44.3	218	1	P25A_BOVIN
20	39	44.3	439	1	AP50_DICDI
21	39	44.3	831	1	FTSR_MYCTU
22	39	44.3	834	1	VW22_YEAST
23	38.5	43.8	857	1	Y2A_CMYNT
24	38	43.2	327	1	FBR1_MOUSE
25	38	43.2	340	1	EBP2_CAEEL
26	38	43.2	370	1	ERWE_SACER
27	38	43.2	476	1	GAB3_CHICK
28	38	43.2	483	1	CLK1_MOUSE
29	38	43.2	512	1	FUS_BOVIN
30	38	43.2	518	1	FUS_MOUSE
31	38	43.2	526	1	FUS_HUMAN
32	38	43.2	636	1	DEDI_SCHPO
33	38	43.2	674	1	TML1_ARATH

ALIGNMENTS

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
AC P20930;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN FIG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89296901; PubMed=2740331;
RA McKinnley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Camitzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
[2]
RP CITRULLINATION.
RX MEDLINE=96374388; PubMed=8780679;
RA Senshu T., Kan S., Ogawa H., Marabe M., Aeaga H.;
RT "Preferential delamination of keratin XI and filaggrin during the
RT terminal differentiation of human epidermis.";
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -I- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -I- PFM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -I- PFM: Undergoes delamination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M24355; AAAS2454.1; -
CC PIR: A32947; A32947.
CC Genew; HGNC:3748; FIG.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.


```

FRQ_CRESP          STANDARD;          PRT; 1015 AA.
ID   FRQ_CRESP
AC   Q00586;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Frequency clock protein.
GN   FRQ.
OS   Crepus spullosus (Chromocera spullosa).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX   NCBI_TaxID=10619;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FGSC 4896;
RX   MEDLINE=97188515; PubMed=90371100;
RA   Lewis M.T., Morgan L.W., Feldman J.F.;
RT   "Analysis of frequency (frq) clock gene homologs: evidence for a
RT   helix-turn-helix transcription factor.";
RL   Mol. Gen. Genet. 253:401-414(1997).
CC   -!- FUNCTION: Circadian clock component involved in the generation of
CC   biological rhythms, in particular in rhythm stability, period
CC   length, and temperature compensation. Behaves as a negative
CC   element in circadian transcriptional loop (by similarity).
CC   -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; U25850; AAA66072.1; -.
DR   FNR; T42013; T42013.
KW   Biological rhythms; Transcription regulation; Nuclear protein.
FT   DOMAIN 240..245      POLY-SER.
FT   DOMAIN 356..368      POLY-GLN.
FT   DOMAIN 443..451      POLY-PRO.
FT   DOMAIN 584..588      NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN 892..913      ASF/GNU-RICH (ACIDIC).
SQ   SEQUENCE 1015 AA; 110972 MW; EA49E732BD7414B1 CRC64;

Query Match          48.9%; Score 43; DB 1; Length 1015;
Best Local Similarity 42.1%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGSGSGS 19
DB 583 NHRKQXTHSTGDSGSSG 601

RESULT 6
NCR_MOUSE          STANDARD;          PRT; 1453 AA.
ID   NCR_MOUSE
AC   P30415;
DT   01-APR-1993 (Rel. 25, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   NK-tumor recognition protein (Natural-killer cells cyclophilin-
DE   related protein) (NK-TR protein).
GN   NKTR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9313824; PubMed=8421688;
RA   Anderson S.R., Gallinger S., Roder J., Frey J., Young H.A.,
RA   Otaido J.R.;
RT   "A cyclophilin-related protein involved in the function of natural

```

```

RT   killer cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RN   [2]
RP   REVISIONS TO C-TERMINUS.
RC   STRAIN=BALB/c; TISSUE=Blood;
RA   Anderson S.R.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC   -!- FUNCTION: Component of a putative tumor-recognition complex.
CC   Involved in the function of NK cells.
CC   -!- SIMILARITY: Contains 1 cyclophilin-like PPIase domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; L04289; AAA37500.2; ALT_INIT.
DR   HSBP; Q27450; IAB3.
DR   MSD; MG197346; NKtr.
DR   InterPro; IPR002130; CSA_PPIase.
DR   Pfam; PF00160; pro_isomerase_1.
DR   PRINTS; PR00153; CSAPISMPASE.
DR   PROSITE; PS00170; CSA_PPIASE_1; 1.
DR   PROSITE; PS50072; CSA_PPIASE_2; 1.
KM   Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
FT   DOMAIN 1..176      PPIASE, CYCLOPHILIN-TYPE.
FT   DOMAIN 222..241      ARG/LYS-RICH (BASIC).
FT   DOMAIN 422..459      ARG/LYS-RICH (BASIC).
FT   DOMAIN 964..1003      ARG/LYS-RICH (BASIC).
FT   DOMAIN 198..273      ARG/SER-RICH.
FT   DOMAIN 468..565      ARG/SER-RICH.
FT   DOMAIN 658..812      ARG/SER-RICH.
FT   DOMAIN 1303..1453    ARG-SER TANDEM REPEAT-RICH.
SQ   SEQUENCE 1453 AA; 163439 MW; DFL173FFB14B283B CRC64;

Query Match          48.9%; Score 43; DB 1; Length 1453;
Best Local Similarity 81.8%; Pred. No. 30;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 RSKGRSGSGSGS 19
DB 682 RSSGSGSGSGS 692

RESULT 7
DEXT_STRDO          STANDARD;          PRT; 1337 AA.
ID   DEXT_STRDO
AC   P39653;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
DE   glucanohydrolase).
GN   DEX.
OS   Streptococcus downei (Streptococcus sobrinus).
OG   Plasmid PYA902.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=13117;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
RC   STRAIN=6715 / UAB65;
RX   MEDLINE=94292401; PubMed=8021165;
RA   Wanda S.-Y., Curtiss R. III;
RT   "Purification and characterization of Streptococcus sobrinus
RT   dextranase produced in recombinant Escherichia coli and sequence
RT   analysis of the dextranase gene.";
RL   J. Bacteriol. 176:3839-3850(1994).
CC   -!- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
CC   PELLICLE-COATED TOOTH SURFACE.

```

```

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
CC linkages in dextran.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
CC and at 39 degrees Celsius.
CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M96978; AAA21772.1; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRfam: TIGR01167; LpxTG_anchor; 1.
DR PROSITE, PS0847; GRAM_POS_ANCHORING; 1.
DR Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
KM Signal; Plasmid.
FT CHAIN 1 30 DEXTRANASE.
FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
FT SITE 1305 1308 LpxTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 1337;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRS 17
Db 1154 ANQDSTKGSSADQSGKS 1170

RESULT 8
MKR2_MOUSE STANDARD; PRT; 774 AA.
ID MKR2_MOUSE
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
DE MARK2 OR EMK.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93364122; PubMed=8358177;
RA Ingalls J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19."
RL Mamm. Genome 4:401-403 (1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X70764; CA50040.1; -.
DR PIR, I48609; I48609.
DR HSSP; Q63450; 1A06.
DR MGD; MGI:99638; Mark2.
DR InterPro: IPR001772; Kinase Cterm.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE, PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT NP_BIND 53 304 PROTEIN_KINASE.
FT BINDING 59 67 ATP (BY SIMILARITY).
FT ACT_SITE 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

Query Match 46.6%; Score 41; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 STYGRSXGRSGRSGRS 19
Db 614 SPFGHSQGRGASGS 628

RESULT 9
SE5_RAT STANDARD; PRT; 825 AA.
ID SE5_RAT
AC Q63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SE5 antigen.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein SE5 in the nervous system."
RL J. Biochem. 118:122-128 (1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D37934; BA07153.1; -.
DR PIR; J04163; J04163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86631 MW; AF667FE2FD55BDF CRC64;

Query Match 46.6%; Score 41; DB 1; Length 825;

```

Best Local Similarity 66.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Cy 8 GRSXGRSGSGS 19
Db 589 GRSXGRSGSGS 600

RESULT 10
COAC YEAST STANDARD; PRT; 2233 AA.
AC 000955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.12) (ACC) [Includes: Biotin
carboxylase (EC 6.3.4.14)].
GN FAS3 OR ACC1 OR YNR016C OR N3175.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycace;
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2015-2022.
RX MEDLINE=92262474; PubMed=1350093;
RA Al-Feel W., Chirala S.S., Wakil S.J.;
RT "Cloning of the yeast FAS3 gene and primary structure of yeast
RT acetyl-CoA carboxylase."
RL Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Poh1 T.M.;
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This protein carries three functions: biotin carboxyl
CC carrier protein, biotin carboxylase, and carboxyltransferase.
CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
CC + malonyl-CoA.
CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -1- COFACTOR: Biotin.
CC -1- ENZYME REGULATION: By phosphorylation.
CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, M2156; AAA20073.1; -;
DR EMBL, Z71631; CA96294.1; -;
DR PIR, S63347; S63347.
DR HSSP, P24182; 1BNC.
DR GenOnline, 143361; -;
DR SGD, S0005299; ACC1.
DR GO, GO:0005789; C:cytoplasmic reticulum membrane; IDA.
DR GO, GO:0003989; F:acetyl-CoA carboxylase activity; IMP.
DR GO, GO:0004075; F:biotin carboxylase activity; IMP.
DR GO, GO:0004075; F:biotin carboxylase activity; IMP.
DR InterPro, IPR001882; Biotin_BS.
DR InterPro, IPR005482; Biotin_carb_C.
DR InterPro, IPR000089; Biotin_lipoyl.
DR InterPro, IPR000022; Carboxyl_trans.
DR InterPro, IPR005479; CPase_L_D2.
DR InterPro, IPR005481; CPase_L_N.
DR Pfam, PF02785; Biotin_carb_C_1.
DR Pfam, PF00364; Biotin_lipoyl_1.
DR Pfam, PF01039; Carboxyl_trans_1.

DR Pfam, PF00289; CPase_L_chain_1.
DR Pfam, PF02786; CPase_L_D2_1.
DR PROSITE, PS00188; BIOTIN; 1.
DR PROSITE, PS00866; CPASE_1; 1.
DR PROSITE, PS00867; CPASE_2; 1.
DR Fatty acid biosynthesis; biotin; ligase; Multifunctional enzyme;
KM ATP-binding; Phosphorylation.
FT NP_BIND 256 261
FT ACT_SITE 383 383
FT BINDING 735 735
FT DOMAIN 1865 1894
FT CONFLICT 1523 1523
FT CONFLICT 1755 1755
FT CONFLICT 1761 1766
SQ SEQUENCE 2233 AA; 250351 MW; 0A335AD9B1F8308 CRC64;
Query Match 46.6%; Score 41; DB 1; Length 2233;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Cy 2 HESTXGRSGSGSGS 19
Db 1200 HSSXNGPAPPGRSGSAS 1217
RESULT 11
DH1_MAIZE STANDARD; PRT; 168 AA.
AC P12950; P16318;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Dehydrin DHN1 (M3) (RAB-17 protein).
GN DHN1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93357436; PubMed=2562763;
RA Close T.J., Kortt A.A., Chandler P.M.;
RT "A cDNA-based comparison of dehydration-induced proteins (dehydrins)
RT in barley and corn."
RL Plant Mol. Biol. 13:95-108 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, AC 1503;
RX MEDLINE=91346629; PubMed=2151715;
RA Villardell J., Goday A., Freire M.A., Torrent M., Martinez M.C.,
RA Torne J.M., Pages M.;
RT "Gene sequence, developmental expression, and protein phosphorylation
RT of RAB-17 in maize."
RL Plant Mol. Biol. 14:423-432 (1990).
RN [3]
RP PHOSPHORYLATION BY CKII.
RX MEDLINE=92042196; PubMed=1939268;
RA Plana M., Iarte E., Ertija R., Goday A., Pages M., Martinez M.C.;
RT "Phosphorylation of maize RAB-17 protein by casein kinase 2."
RL J. Biol. Chem. 266:22510-22514 (1991).
CC -1- DEVELOPMENTAL STAGE: ABA treatment induced the synthesis of RAB-17
CC in calli, however, the RAB-17 proteins were found to be highly
CC phosphorylated only in embryos.
CC -1- INDUCTION: By abscisic acid (ABA) and water stress.
CC -1- PTM: SERINE IS THE ONLY AMINO ACID WHICH BECOMES PHOSPHORYLATED.
CC -1- SIMILARITY: Belongs to the plant dehydrin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL, X15290; CAA33364.1; -
 DR EMBL, X15994; CAA34123.1; -
 DR PIR, A39316; A39316.
 DR PIR, S05545; S05545.
 DR PIR, S08633; S08633.
 DR Malzeda, 24916; -
 DR InterPro; IPR00167; Dehydrin.
 DR Pfam; PF00257; dehydrin.1.
 DR ProSite; PS00315; DEHYDRIN_1; 1.
 DR ProSite; PS00823; DEHYDRIN_2; 2.
 DR Dehydrin; Repeat; Phosphorylation; Multigene family.
 KW DOMAIN 76 84 POLY-SER
 FT DOMAIN 2 140 3 X APPROXIMATE REPEATS.
 FT REPEAT 2 10 1-1.
 FT REPEAT 122 130 1-1.
 FT REPEAT 132 140 1-2.
 FT DOMAIN 94 166 1-3.
 FT REPEAT 94 107 2 X APPROXIMATE REPEATS.
 FT REPEAT 153 166 2-1.
 FT REPEAT 9 9 2-2.
 FT CONFLICT 9 9 R -> H (IN REF. 2).
 FT CONFLICT 13 13 R -> A (IN REF. 2).
 FT CONFLICT 41 41 MISSING (IN REF. 2).
 SQ SEQUENCE 168 AA; 17161 MW; BE7758CAD37CF39 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 168;
 Best Local Similarity 38.9%; Pred. No. 9.3;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 HOESTXGRSGRSGSGS 19
 DB 114 HATATTGGAYGGGHTGS 131

RESULT 12
 SOL2_YEAST STANDARD; PRT; 315 AA.
 AC P37262;
 DT 01-OCT-1996 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (6PGL).
 GN SOL2 OR YCRX13W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96363906; PubMed=8725220;
 RA Shen W.C., Stanford D.R., Hopper A.K.;
 RT "Loss of the 6-phosphogluconolactonase gene, positively regulates
 RT members of the SOL gene family.";
 RL Genetics 143:699-712(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ballester J.P.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
 RA Sanz E.;
 RT Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION AND SIMILARITY.
 RX MEDLINE=94147996; PubMed=8313894;
 RA Koonin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RL EMBL J. 13:493-503(1994).
 CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
 CC phosphogluconate. (By similarity).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
 CC phospho-D-gluconate.
 CC -1- PATHWAY: Pentose phosphate pathway; second step.
 CC -1- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate

CC isomerase family. 6-phosphogluconolactonase subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL, U46559; AAB49322.1; -
 DR EMBL, X59720; CAA42272.1; -
 DR PIR, S53589; S53589.
 DR Germonline; 139004; -
 DR SGD; S0000718; SOL2.
 DR GO; GO:0008033; P:RNA processing; IGI.
 DR InterPro; IPR006146; Gluc_gal_isom.
 DR InterPro; IPR005900; Phosphogluconlac.
 DR Pfam; PF01182; Glucosamine_iso; 1.
 DR TIGRFAMs; TIGR01198; pgl; 1.
 KW Hydrolyase.
 FT CONFLICT 171 171 A -> V (IN REF. 2).
 SQ SEQUENCE 315 AA; 34501 MW; 7F0AAD76574AB276 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 315;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 4 ESTXGRSGRSGSGS 19
 DB 60 KSTSAAEGRSGSGS 75

RESULT 13
 CLTK1_HUMAN STANDARD; PRT; 484 AA.
 AC P49759;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dual specificity protein kinase CLK1 (EC 2.7.1.37) (EC 2.7.1.112) (CDC
 DE dual kinase 1).
 GN CLTK1 OR CLK.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91138618; PubMed=1704859;
 RA Johnson K.W., Smith K.A.;
 RT "Molecular cloning of a novel human cdc2/CDC28-like protein kinase.";
 RL J. Biol. Chem. 266:3402-3407(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95082033; PubMed=7990150;
 RA Hanes J.J., der Kammer H., Klauting J.J., Scheit K.H.;
 RT "Characterization by cDNA cloning of two new human protein kinases.
 RT Evidence by sequence comparison of a new family of mammalian protein
 RT kinases.";
 RL J. Mol. Biol. 244:665-672(1994).
 CC -1- FUNCTION: PHOSPHORYLATES SERINE- AND ARGinine-RICH (SR) PROTEINS
 CC OF THE SPliceosomal COMPLEX MAY BE A CONSTITUENT OF A NETWORK OF
 CC REGULATORY MECHANISMS THAT ENABLE SR PROTEINS TO CONTROL RNA
 CC SPLICING. PHOSPHORYLATES SERINES, THREONINES AND TYROSINES (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: Named isoforms=2;
 CC Name=Long;

```
CC      IsoId=P49759-1; Sequence=Displayed;
CC      Name=Short;
CC      IsoId=P49759-2; Sequence=VSP_004852, VSP_004853;
CC      Note=Lacks the kinase domain;
CC      -1- P1M: Autophosphorylates on all three types of residues (By
CC      similarity). Belongs to the Ser/Thr family of protein kinases.
CC      Lammer subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; I29219; AAA61480.1; -
DR      EMBL; I29222; AAB59459.1; -
DR      PIR; S53641; S53641.
DR      HSSP; Q00534; 1B18.
DR      Genew; HGNC:2068; CLK1.
DR      GK; P49759; -
DR      MM; 601951; -
DR      GO; GO:0004715; P:non-membrane spanning protein tyrosine kina. . .; TAS.
DR      GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR      GO; GO:0008283; P:cell proliferation; TAS.
DR      GO; GO:0000074; P:regulation of cell cycle; TAS.
DR      InterPro; IPR000719; Prot_Kinase.
DR      InterPro; IPR008271; Ser_Thr_pkin_AS.
DR      InterPro; IPR002290; Ser_Thr_pkinase.
DR      Pfam; PF00069; Pkinase; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SMO0220; S_TKc; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      TranserBase; Serine/threonine-protein kinase; ATP-binding;
DR      Tyrosine-protein kinase; Phosphorylation; Nuclear protein;
DR      Alternative splicing.
DR      DOMAIN 161 477 PROTEIN KINASE.
DR      NP_BIND 167 175 ATP (BY SIMILARITY).
DR      BINDING 191 191 ATP (BY SIMILARITY).
DR      ACT_SITE 288 288 BY SIMILARITY.
DR      VARSPLIC 131 136 KSHRRK -> MKLLIL (in isoform Short).
DR      FT /FTID=VSP_004852.
DR      FT VARSPLIC 137 484 Missing (in isoform Short).
DR      FT /FTID=VSP_004853.
DR      FT SEQUENCE 484 AA; 57205 MM; 3045B9486AD0A6B CRC64;
DR      Query Match 45.5%; Score 40; DB 1; Length 484;
DR      Best Local Similarity 47.1%; Pred. No. 29;
DR      Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 SHOESTXGRSGRSGRS 17
DB      95 SRYQNHSSKSGRSGRS 111

RESULT 14
ZABA YEAST STANDARD; PRT; 526 AA.
AC      Q00362;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Protein phosphatase PP2A regulatory subunit B (PR55) (Cell division
DE      control protein 55).
GN      CDC55 OR YGL190C OR G1345.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX      NCBI_TaxId=4932;
```

```
RM      [1]
RM      SEQUENCE FROM N.A.
RM      MEDLINE=92017858; PubMed=1656238;
RM      Healy A.M., Zolnerwicz S., Stapleton A.E., Goebel M.,
RM      Depauli-Roch A.A., Pringle J.R.;
RM      "CDC55, a Saccharomyces cerevisiae gene involved in cellular
RM      morphogenesis: identification, characterization, and homology to the
RM      B subunit of mammalian type 2A protein phosphatase.";
RM      Mol. Cell. Biol. 11:5767-5780(1991).
RM      [2]
RM      SEQUENCE FROM N.A.
RM      STRAIN=5288c / FY1679;
RM      MEDLINE=97197971; PubMed=9046087;
RM      Coglievina M., Klima R., Berani I., Delneri D., Zaccaria P.,
RM      Brusch C.V.;
RM      "Sequencing of a 40.5 kb fragment located on the left arm of
RM      chromosome VII from Saccharomyces cerevisiae.";
RM      Yeast 13:55-64(1997).
RM      [3]
RM      FUNCTION: Phosphatase 2A affects a variety of biological processes
RM      in the cell such as transcription, cell cycle progression and
RM      cellular morphogenesis, and provides an initial identification of
RM      critical substrates for this phosphatase. The regulatory subunit
RM      may direct the catalytic subunit to distinct, albeit overlapping,
RM      subsets of substrates.
RM      [4]
RM      SUBUNIT: PP2A exists in several trimeric forms, all of which
RM      consist of a core composed of a catalytic subunit associated with
RM      a 65 kDa (PR65) (Subunit A) and a 55 kDa (PR55) (Subunit B)
RM      regulatory subunit.
RM      [5]
RM      SIMILARITY: Belongs to the phosphatase 2A regulatory subunit B
RM      family.
RM      [6]
RM      -1- SIMILARITY: Contains 7 WD repeats.
RM      [7]
RM      This SWISS-PROT entry is copyright. It is produced through a collaboration
RM      between the Swiss Institute of Bioinformatics and the EMBL outstation -
RM      the European Bioinformatics Institute. There are no restrictions on its
RM      use by non-profit institutions as long as its content is in no way
RM      modified and this statement is not removed. Usage by and for commercial
RM      entities requires a license agreement (See http://www.isb-sib.ch/announce/
RM      or send an email to license@isb-sib.ch).
RM      -----
DR      EMBL; M72716; AAA34482.1; -
DR      EMBL; Z72712; CAA96902.1; -
DR      EMBL; X91837; CAA62954.1; -
DR      EMBL; X91489; CAA62785.1; -
DR      PIR; A41698; A41698.
DR      Genmonline; 141238; -
DR      SGD; S0003158; CDC55.
DR      GO; GO:0007094; P:mitotic spindle checkpoint; IMP.
DR      GO; GO:0006412; P:protein biosynthesis; IMP.
DR      GO; GO:0007124; P:pseudohyphal growth; IMP.
DR      InterPro; IPR000009; PP2A_PR55.
DR      InterPro; IPR001680; WD40.
DR      Pfam; PF00400; WD40; 4.
DR      PRINTS; PR00600; PP2APR55.
DR      SMART; SMO0320; WD40; 3.
DR      PROSITE; PS01024; PR55-1; 1.
DR      PROSITE; PS01025; PR55-2; 1.
DR      PROSITE; PS00678; WD_REPEATS_1; 1.
DR      PROSITE; PS50082; WD_REPEATS_2; FALSE NEG.
DR      PROSITE; PS50294; WD_REPEATS_REGION; FALSE NEG.
DR      Repeat: WD repeat; Cell cycle.
DR      REPEAT 82 123 WD 1.
DR      REPEAT 82 123 WD 1.
DR      REPEAT 182 220 WD 2.
DR      REPEAT 231 271 WD 4.
DR      REPEAT 290 328 WD 5.
DR      REPEAT 345 386 WD 6.
DR      REPEAT 495 525 WD 7.
DR      DOMAIN 416 419 POLY-SER.
DR      FT CONFLICT 500 500 I -> N (IN REF. 1).
DR      FT SEQUENCE 526 AA; 59662 MM; 6DA12C28057A6A82 CRC64;
DR      Query Match 45.5%; Score 40; DB 1; Length 526;
```


Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-7

Perfect score: 88

Sequence: 1 SHQSTYGRSGXGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	85	96.6	798	4 Q9H4U3	Q9H4U3 homo sapien
2	85	96.6	1084	4 Q01212	Q01212 homo sapien
3	84	95.5	797	4 Q16824	Q16824 homo sapien
4	84	95.5	980	4 Q15206	Q15206 homo sapien
5	84	95.5	1218	4 Q05531	Q05531 homo sapien
6	76	86.4	591	4 Q01720	Q01720 homo sapien
7	76	86.4	687	4 Q9H4U2	Q9H4U2 homo sapien
8	71	80.7	465	4 Q03838	Q03838 homo sapien
9	70	79.5	322	4 Q75370	Q75370 homo sapien
10	51	58.0	1284	5 Q9V9Y3	Q9V9Y3 drosophila
11	49	55.7	822	3 Q9P312	Q9P312 neurospora
12	48	54.5	471	5 Q9UAY0	Q9UAY0 caenorhabdit
13	47	53.4	306	5 Q45386	Q45386 caenorhabdit
14	47	53.4	399	4 Q9NTA9	Q9NTA9 homo sapien
15	47	53.4	436	16 Q82C67	Q82C67 streptomyces
16	47	53.4	467	4 Q9BW18	Q9BW18 homo sapien

17	47	53.4	737	4 Q9BQ39	Q9BQ39 homo sapien
18	46	52.3	233	10 Q9AX45	Q9AX45 cryza serrat
19	46	52.3	426	4 Q8WV76	Q8WV76 homo sapien
20	45	51.1	259	4 Q86YA2	Q86YA2 homo sapien
21	45	51.1	669	11 Q8BRU5	Q8BRU5 mus musculus
22	45	51.1	734	11 Q9NM79	Q9NM79 mus musculus
23	44	50.0	216	4 Q9H711	Q9H711 homo sapien
24	44	50.0	322	16 Q8G4F5	Q8G4F5 bifidobacter
25	44	50.0	306	5 Q15845	Q15845 leishmania
26	44	50.0	936	16 Q82H71	Q82H71 streptomyces
27	43.5	49.4	286	13 Q91988	Q91988 xiphophorus
28	43	48.9	446	16 Q9AX30	Q9AX30 streptomyces
29	43	48.9	649	16 Q83749	Q83749 treponema p
30	43	48.9	813	12 Q91IN8	Q91IN8 hepatitis b
31	43	48.9	820	12 Q68375	Q68375 mouse cytom
32	43	48.9	869	10 Q9C7M2	Q9C7M2 arabidopsis
33	43	48.9	849	10 Q949M6	Q949M6 arabidopsis
34	43	48.9	966	11 Q8CBP6	Q8CBP6 mus musculus
35	43	48.9	1064	3 Q96X23	Q96X23 yarrowia li
36	43	48.9	1142	13 Q7S2V2	Q7S2V2 xenopus lae
37	42.5	48.3	240	16 Q87YL4	Q87YL4 pseudomonas
38	42	47.7	297	6 Q8SQ24	Q8SQ24 bos taurus
39	42	47.7	481	16 Q7UDP9	Q7UDP9 rhodospirillum
40	42	47.7	552	11 Q8K028	Q8K028 mus musculus
41	42	47.7	658	11 Q8C5U9	Q8C5U9 mus musculus
42	42	47.7	806	12 Q91522	Q91522 hepatitis b
43	42	47.7	809	5 Q96OC5	Q96OC5 drosophila
44	42	47.7	1135	5 Q9VJB6	Q9VJB6 drosophila
45	42	47.7	1229	3 Q86ZM1	Q86ZM1 neurospora

ALIGNMENTS

RESULT 1	ID	Q9H4U3	PRELIMINARY;	PRT;	798 AA.
AC	Q9H4U3				
DT	01-MAR-2001 (TRENBLREL.16, Created)				
DT	01-MAR-2001 (TRENBLREL.16, Last sequence update)				
DT	01-JUN-2003 (TRENBLREL.24, Last annotation update)				
DE	DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).				
GN	PLG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Latid G.;				
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AL356504; CAC13171.1; -				
DR	GO; GO:0005198; F:structural molecule activity; IEA.				
DR	InterPro; IPR003303; Filaggrin.				
DR	Pfam; PF03516; Filaggrin; 4.				
DR	PRINTS; PR00487; FILAGGRIN.				
FT	NON_TER				
FT	SEQUENCE	798 AA;	84773 MW;	F923DDABD1290805 CRC64;	
QY	Query Match	96.6%;	Score 85;	DB 4;	Length 798;
DB	Best Local Similarity.	89.5%;	Pred. No. 5e-06;		
Matches	17;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
QY	1 SHQSTYGRSGXGRSGSGS 19				
DB	428 SHQSTYGRSGXGRSGSGS 446				
RESULT 2					
ID	Q01212	PRELIMINARY;	PRT;	1084 AA.	
AC	Q01212; Q03840;				
DT	01-NOV-1996 (TRENBLREL.01, Created)				

```
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-UN-2003 (TrEMBLrel. 24, last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER
SQ SEQUENCE 1084 AA; 115271 MW; 80046408BD5A362D CRC64;

Query Match 96.5%; Score 85; DB 4; Length 1084;
Best Local Similarity 89.5%; Pred. No. 7e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGRSGS 19
Db 64 SHQSTGRSGRSGRSGS 82

RESULT 3
Q16824 PRELIMINARY; PRT; 797 AA.
ID Q16824;
AC Q16824;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60502; AAA63248.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER
SQ SEQUENCE 797 AA; 85176 MW; 60E184763BD86B CRC64;

Query Match 95.5%; Score 84; DB 4; Length 797;
Best Local Similarity 89.5%; Pred. No. 7.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGRSGS 19
Db 427 SHQSTGRSGRSGRSGS 445
```

```
RESULT 4
Q15206 PRELIMINARY; PRT; 990 AA.
ID Q15206;
AC Q15206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-UN-2003 (TrEMBLrel. 24, last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11:30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60494; AAA63244.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; FILAGGRIN.
FT NON TER
SQ SEQUENCE 990 AA; 106453 MW; A8396F10FA91991 CRC64;

Query Match 95.5%; Score 84; DB 4; Length 990;
Best Local Similarity 89.5%; Pred. No. 9.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGRSGS 19
Db 875 SHQSTGRSGRSGRSGS 893

RESULT 5
Q05331 PRELIMINARY; PRT; 1218 AA.
ID Q05331;
AC Q05331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
```

324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

-1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.

EMBL; M96943; AAA36487.1; -.

DR PIR; A48118; A48118.

DR HSP; P02593; 1CDM.

DR GO; GO:0005856; C:cytoskeleton; NAS.

DR GO; GO:0005509; F:calcium ion binding; TAS.

DR GO; GO:0030154; P:cell differentiation; NAS.

DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.

DR InterPro; IPR001751; CABP_S100.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF03516; Filaggrin; 6.

DR Pfam; PF01023; S_100; 1.

DR PRINTS; PR00487; FILAGGRIN.

DR PROSITE; PS00018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CABP; 1.

KW Phosphorylation; Polyprotein; Developmental protein; Calcium-binding; Polymorphism.

KW CA_BIND 19 32 SITE I (BY SIMILARITY).

FT CA_BIND 62 73 SITE II (BY SIMILARITY).

FT NON_TER 1218 1218

SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match Best Local Similarity 95.5%; Score 84; DB 4; Length 1218; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 89.5%; Pred. No. 1.2e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 SHQESTXGRSGRSGSGS 19

DB 1097 SHQESTXGRSGRSGSGS 1115

RESULT 6

ID Q01720 PRELIMINARY; PRT; 591 AA.

AC Q01720;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN (1) _TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=93054736; PubMed=1429717;

RA Presland R.B., Haydock P.V., Fleckman P., Ntunukisiri W., Dale B.A.; "Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus"; J. Biol. Chem. 267:23772-23781 (1992).

RL J. Biol. Chem. 267:23772-23781 (1992).

CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.

EMBL; L01089; AAA60177.1; -.

DR EMBL; L01090; AAA60176.1; -.

DR PIR; A45135; A45135.

DR PIR; A48118; A48118.

DR HSP; P80511; 1B8A.

DR MIM; 135940; -.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR001751; CABP_S100.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF03516; Filaggrin; 2.

DR Pfam; PF01023; S_100; 1.

DR PRINTS; PR00487; FILAGGRIN.

DR PROSITE; PS00018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CABP; 1.

KW Polymorphism.

FT PROPEP 1 293 POTENTIAL.

FT CHAIN 294 467 FILAGGRIN.

FT PROPEP 468 474 POTENTIAL.

FT CHAIN 475 >591 FILAGGRIN.

FT CA_BIND 19 32 SITE I (BY SIMILARITY).

FT CA_BIND 62 73 SITE II (BY SIMILARITY).

FT NON_TER 591 591

SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match Best Local Similarity 86.4%; Score 76; DB 4; Length 591; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Best Local Similarity 84.2%; Pred. No. 0.00013;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 SHQESTXGRSGRSGSGS 19

DB 449 SHQESTXGRSGRSGSGS 467

RESULT 7

ID Q9H4U2 PRELIMINARY; PRT; 687 AA.

AC Q9H4U2;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE DJ14N1.1.1 (profilaggrin 5' end) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN (1) _TaxID=9606;

RP SEQUENCE FROM N.A.

RA Laird G.;

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

EMBL; AL356504; CAC13172.1; -.

DR PIR; A48118; A48118.

DR HSP; P80511; 1B8A.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR001751; CABP_S100.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF03516; Filaggrin; 3.

DR Pfam; PF01023; S_100; 1.

DR PRINTS; PR00487; FILAGGRIN.

DR PROSITE; PS00018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CABP; 1.

FT NON_TER 687 687

SQ SEQUENCE 687 AA; 76659 MW; 8000363BEBF07B74 CRC64;

Query Match Best Local Similarity 86.4%; Score 76; DB 4; Length 687; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 449 SHOESTXGRSGRSGSGS 467

RESULT 8
 ID 003838 PRELIMINARY; PRT; 465 AA.
 AC 003838:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE FILAGRIN (PROFILAGRIN) (Fragment).
 GN FLG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 29:9432-9440(1990).
 RT [2]
 RP REVISIONS.
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene.";
 RL Biochemistry 30:5814-5814(1991).
 CC -1- FUNCTION: FILAGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
 PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
 FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGRIN
 CC REPEATS.
 CC -1- MISCELLANEOUS: FILAGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
 CC INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDDEM
 CC COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN
 CC GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 CC PROTEOLYTICALLY CLEAVED.
 DR EMBL, M60499; AAA63246.1; -;
 DR GO:GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR003303; Filaggrin.
 DR Pfam: PF03516; Filaggrin; 3.
 DR PRINTS: PR00487; FILAGRIN.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 465 AA; 50280 MW; C883744C5B134097 CRC64;

Query Match 80.7%; Score 71; DB 4; Length 465;
 Best Local Similarity 78.9%; Pred. No. 0.00069;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 227 SHOESTXGRSGRSGSGS 245

RESULT 9
 ID 075370 PRELIMINARY; PRT; 322 AA.
 AC 075370:
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Bidermal filaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99101527; PubMed=9886436;
 RA Girbal-Neuhauser E., Durieux J.J., Arnaud M., Dalbon P., Sabbag M.,
 RA Vincent C., Simon M., Simon T., Masson-Besleiere C.,
 RA Jolivet-Reynaud C., Jolivet M., Serre G.;
 RT "The epitopes targeted by the rheumatoid arthritis-associated
 RT anti-filaggrin autoantibodies are posttranslational generated on
 RT various sites of (pro)filaggrin by deamination of arginine residues.";
 RL J. Immunol. 162:585-594(1999).
 DR EMBL, AF043380; AAC23559.1; -;
 DR GO:GO:0005198; F:structural molecule activity; IEA.
 DR InterPro: IPR003303; Filaggrin.
 DR Pfam: PF03516; Filaggrin; 2.
 DR PRINTS: PR00487; FILAGRIN.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF950 CRC64;

Query Match 79.5%; Score 70; DB 4; Length 322;
 Best Local Similarity 77.8%; Pred. No. 0.00068;
 Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 18
 DB 305 SHOESTXGRSGRSGSGS 322

RESULT 10
 ID 09V9Y3 PRELIMINARY; PRT; 1284 AA.
 AC 09V9Y3:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE CG11339-2A.
 GN CG11339.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer V.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abritil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernson B.F., Bhandari D., Bolshakov S.,
 RA Bokoyeva D., Borkan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibbegan C.,
 RA Urali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reibert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-lamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Mathews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Bernier B.P., Carlson J.W., Celisner S.E.,
 RA Clump M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.U.,
 RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
 RA Whitfield E.U., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.,
 RT "Annotation of *Drosophila melanogaster* genome,"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AEO03777; AAF57145.2; -
 DR Flybase; FBgn0039841; CG11339.
 DR GO; GO:0005856; C:cytoskeleton; IEA.
 DR InterPro; IPR000299; Band 4.1.
 DR Pfam; PF00373; Band 4.1; 1.
 DR PRINTS; PR00935; BAND41.
 DR SMART; SM00295; B41; 1.
 DR PROSITE; PS50057; FERM_3; 1.
 SQ SEQUENCE 1284 AA; 138966 MW; EDF4119838AA6447 CRC64;
 SO Query Match 58.0%; Score 51; DB 5; Length 1284;
 Best Local Similarity 61.1%; Pred. No. 5.8;
 Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGS 18
 Db 545 SNESELSRSGRSGRSG 562
 RESULT 11
 QP312 PRELIMINARY; PRT; 822 AA.
 AC QP312;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DS Related to nuclear phosphoprotein.
 GN B12P1.10.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_Taxid=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL390091; CAB98213.1; -.

DR PIR; T51049; T51049.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR005120; Smg-4_UPF3.
 DR Pfam; PF00076; trm; 1.
 DR Pfam; PF03467; Smg4_UPF3; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;
 SO Query Match 55.7%; Score 49; DB 3; Length 822;
 Best Local Similarity 52.9%; Pred. No. 7.9;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 QESTXGRSGRSGS 19
 Db 414 RESASGRTRGRGRGCT 430
 RESULT 12
 Q9UAY0 PRELIMINARY; PRT; 471 AA.
 AC Q9UAY0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE W03G1.5 protein.
 GN W03G1.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson R., Baynes C., Barks M.,
 RA Bonfield J., Burton U., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Crixton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkes T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kistler J., Laister N., Latreille P.,
 RA Latching U., Lloyd C., McMurray A., Northmore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rikken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstein L., Wilkinson-Spratt J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*,"
 RL Nature 368:32-38 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Pauley A., Scheet P., Harper M.,
 RT "The sequence of *C. elegans* cosmid W03G1,"
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.,
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF125964; AAD14753.1; -.
 DR PIR; T33997; T33997.
 DR WormPep; W03G1.5; CE17283.
 SQ SEQUENCE 471 AA; 50885 MW; BDF30B59A64A955B CRC64;
 SO Query Match 54.5%; Score 48; DB 5; Length 471;
 Best Local Similarity 52.6%; Pred. No. 6.3;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGS 19
 Db 167 SSRSPSRGRGRGRGRSGS 185

RESULT 13

045386 PRELIMINARY; PRT; 306 AA.
 AC 045386;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE F21H7.5 protein.
 OS F21H7.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gardner A.E.;
 RN Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RX SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "genome sequence of the nematode C. elegans: A platform for
 investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z93379; CAB07388.1; -.
 DR PIR; T21220; T21220.
 DR WormPep; F21H7.5; C158877.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000535; MSP domain.
 DR InterPro; IPR008962; Papd-like.
 DR Pfam; PF006635; MSP domain; 1.
 DR PROSITE; PS50202; MSP; 1.
 SQ SEQUENCE 306 AA; 32421 MW; 830C17DB321DDCC28 CRC64;

Query Match 53.4%; Score 47; DB 5; Length 306;
 Best Local Similarity 44.4%; Pred. No. 5.8;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSG 18
 Db 48 SHRSSKTKRSGRSGSG 65

RESULT 14
 Q9NTA9 PRELIMINARY; PRT; 399 AA.
 AC Q9NTA9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS DKFZP761E0323.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bioecker H., Boecker M., Brandt P., Mewes H.W., Gassenhuber J.,
 RA Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137423; CAB70733.1; -.
 DR PIR; T46259; T46259.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; Helicase_C; 1.

DR SMART; SMO0490; HELIC; 1.
 KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.
 FT NON TER 1
 SQ SEQUENCE 399 AA; 44075 MW; BC09B3A0E4CE5B96 CRC64;

Query Match 53.4%; Score 47; DB 4; Length 399;
 Best Local Similarity 55.6%; Pred. No. 7.7; Indels 0; Gaps 0;
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSG 18
 Db 336 SNRQRSGWSSGRSGRSG 353

RESULT 15

082C67 PRELIMINARY; PRT; 436 AA.
 AC 082C67;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative integral membrane protein.
 GN SAV5487.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005043; BAC73199.1; -.
 DR InterPro; IPR002194; Chaperonin_TCP-1.
 DR PROSITE; PS00995; TCP1_3; 1.
 DR Complete proteome.
 SQ SEQUENCE 436 AA; 45359 MW; 35BD46B80C7282D0 CRC64;

Query Match 53.4%; Score 47; DB 16; Length 436;
 Best Local Similarity 83.3%; Pred. No. 8.5;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GRSXGRSGRSGS 19
 Db 253 GRSXGRSGRSGS 264

Search completed: September 28, 2004, 06:12:46
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQSTXKXSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of resules predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	AAW61515	AAW61515 Peptide c
2	84	95.5	19	AAW61507	AAW61507 Peptide c
3	84	95.5	19	AAW61517	AAW61517 Peptide c
4	84	95.5	19	AAW61511	AAW61511 Peptide c
5	84	95.5	19	AAW61505	AAW61505 Peptide c
6	84	95.5	19	AAW61514	AAW61514 Peptide c
7	84	95.5	19	AAW61516	AAW61516 Peptide c
8	84	95.5	19	AAW61516	AAW61516 Peptide c
9	84	95.5	19	AAW61516	AAW61516 Peptide c
10	84	95.5	19	AAW61516	AAW61516 Peptide c
11	84	95.5	19	AAW61512	AAW61512 Peptide c
12	84	95.5	19	AAW61509	AAW61509 Peptide c
13	84	95.5	19	AAW61513	AAW61513 Peptide c
14	84	95.5	19	AAW61510	AAW61510 Peptide c
15	84	95.5	19	AAW61520	AAW61520 Peptide c
16	84	95.5	19	AAW61520	AAW61520 Peptide c
17	84	95.5	19	AAW61520	AAW61520 Peptide c
18	84	95.5	19	AAW61520	AAW61520 Peptide c
19	84	95.5	19	AAW61520	AAW61520 Peptide c
20	84	95.5	19	AAW61520	AAW61520 Peptide c
21	84	95.5	19	AAW61520	AAW61520 Peptide c
22	84	95.5	19	AAW61520	AAW61520 Peptide c
23	84	95.5	19	AAW61520	AAW61520 Peptide c
24	84	95.5	19	AAW61520	AAW61520 Peptide c
25	84	95.5	19	AAW61520	AAW61520 Peptide c

26	47	53.4	574	4	AAW41855	AAW41855 Human pol
27	47	53.4	752	4	AAU23535	AAU23535 Novel hum
28	44	50.0	106	3	AAQ33285	AAQ33285 Zee mays
29	44	50.0	108	3	AAQ33268	AAQ33268 Zee mays
30	44	50.0	138	4	AAU50949	AAU50949 Propionib
31	44	50.0	138	6	ABM47468	ABM47468 Propionib
32	44	50.0	233	3	AAQ33195	AAQ33195 Zee mays
33	43.5	49.4	286	2	AAW54411	AAW54411 Neurotrop
34	43.5	49.4	19938	6	ABP76679	ABP76679 Streptomy
35	43	48.9	92	4	AAU60843	AAU60843 Propionib
36	43	48.9	92	6	ABM57362	ABM57362 Propionib
37	43	48.9	477	4	ABG08221	ABG08221 Novel hum
38	43	48.9	641	4	ABG19110	ABG19110 Novel hum
39	43	48.9	649	6	ABU48672	ABU48672 Protein e
40	43	48.9	694	2	AAW1267	AAW1267 Drosophi1
41	43	48.9	694	4	ABW1797	ABW1797 Drosophi1
42	43	48.9	982	2	AAW13320	AAW13320 Murine Na
43	43	48.9	982	6	ABU05134	ABU05134 Human exp
44	43	48.9	1482	7	ADW58671	ADW58671 Rat Prote
45	43	48.9	1482	7	ADW58675	ADW58675 Rat Prote

ALIGNMENTS

RESULT 1
AAW61515
ID AAW61515 standard; peptide; 19 AA.
XX
AC AAW61515;
XX
DT 26-OCT-1998 (first entry)
XX
DE Peptide cfa, based on cDNA of a profilaggrin repeat.
XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
RV solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
PV WC9822503-A2.
XX
PD 28-MAY-1998.
XX
PE 14-NOV-1997; 97WC-ND000624.
XX
PR 15-NOV-1996; 96NL-01004539.
XX
PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPE.
XX
PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
DR
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
PS Disclosure; Page 6; 19PP; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
SQ Sequence 19 AA;

Query Match 96.6%; Score 85; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 2

AAW61507 ID AAW61507 standard; peptide; 19 AA.
XX AAW61507;
AC AAW61507;
XX 26-OCT-1998 (first entry)
XX
XX
XX Peptide cfc3, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 11 /note="Citruiline"
XX
XX W09822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97MO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 94.7%; Pred. No. 1.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 3

AAW61517 ID AAW61517 standard; peptide; 19 AA.
XX AAW61517;
AC AAW61517;
XX 26-OCT-1998 (first entry)
XX
XX
XX Peptide cfcQ, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX Homo sapiens.
XX
XX W09822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97MO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 4

AAW61511 ID AAW61511 standard; peptide; 19 AA.
XX AAW61511;
AC AAW61511;
XX 26-OCT-1998 (first entry)
XX
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX Homo sapiens.
XX

```

XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
FT Modified-site 11 /note= "Citruilline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-06;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SHOESTXGRSGRSGS 19
XX 1 SHOESTXGRSGRSGS 19
XX
XX RESULT 5
XX AAW61505
XX ID AAW61505 standard; peptide; 19 AA.
XX AAW61505;
XX 26-OCT-1998 (first entry)
XX Peptide cfcl, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruilline"
FT Modified-site 7 /note= "Citruilline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX

```

```

PF 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA:
SQ
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 1.2e-06;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 SHOESTXGRSGRSGS 19
XX 1 SHOESTXGRSGRSGS 19
XX
XX RESULT 6
XX AAW61514
XX ID AAW61514 standard; peptide; 19 AA.
XX AAW61514;
XX 26-OCT-1998 (first entry)
XX Peptide cf, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX Homo sapiens.
XX
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEME-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA,
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX

```

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 1.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19

DB 1 SHOESTRGRSGRSGSGS 19

RESULT 7

ID AAM61516 standard; peptide; 19 AA.

XX

AC AAM61516;

XX

DT 26-OCT-1998 (first entry)

XX

DE Peptide cFE, based on cDNA of a profilaggrin repeat.

XX

KW Antigen, autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KW monoclonal antibody.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9822503-A2.

XX

PD 28-MAY-1998.

XX

PF 14-NOV-1997; 97WO-NL000624.

XX

PR 15-NOV-1996; 96NL-01004539.

XX

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX

PI (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX

PI Van Venrooij WTM, Schellekens GA, Raats JWH, Host RMA;

DR WPI; 1998-398613/34.

XX

PT Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 1.2e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19

DB 1 SHOESTRGRSGRSGSGS 19

RESULT 8

ID ABB97605 standard; protein; 1467 AA.

XX

AC ABB97605;

XX

DT 27-JUN-2002 (first entry)

XX

DE Novel human protein SEQ ID NO: 873.

XX

KW Human; anti-neuritic; vulnery; anti-inflammatory; immunomodulator;

KW anti-infectivity; cerebroprotective; cytosolic; rheumatic; gene therapy;

KW neuroprotective; antiparkinsonian; protein therapy; Est;

XX

OS Homo sapiens.

XX

PN WO200222660-A2.

XX

PD 21-MAR-2002.

XX

PF 10-SEP-2001; 2001WO-US026015.

XX

PR 11-SEP-2000; 2000US-00659671.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX

DR WPI; 2002-292408/33.

XX

DR N-PSDB; ABN32791.

XX

PT An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

XX

PS Example 2; SEQ ID NO 873; 503pp; English.

XX

CC The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

XX

SQ Sequence 1467 AA;

Query Match 95.5%; Score 84; DB 5; Length 1467;

Best Local Similarity 89.5%; Pred. No. 7.3e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19

DB 1097 SHOESTRGRSGRSGSGS 1115

RESULT 9

ID AAM61506 standard; peptide; 19 AA.

XX

AC AAM61506;

XX

DT 26-OCT-1998 (first entry)
 XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX
 XX Disclosure; Page 6; 19pp; English.
 PS
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 DB 1 SHQESTXGRSGRSGSGS 19
 RESULT 10
 AAW61508
 ID AAW61508 standard; peptide; 19 AA.
 XX
 AC AAW61508;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc4, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 13 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 XX Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 XX
 XX Disclosure; Page 6; 19pp; English.
 PS
 XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSGRSGSGS 19
 DB 1 SHQESTXGRSGRSGSGS 19
 RESULT 11
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 AC AAW61512;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT
 XX
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX

PR	15-NOV-1996;	96NL-01004539.
XX	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
PI	Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;	
XX	WPI; 1998-398613/34.	
DR		
PT	Peptide derived from an antigen recognised by autoantibodies - is	
FT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
PS	used in diagnosis of the disease.	
XX		
PS	Disclosure; Page 6, 19pp; English.	
CC		
CC	Sequences AAM61505-W61520 are peptides derived from the C-terminal end of	
CC	the profilaggrin antigen which is recognised by autoantibodies from	
CC	patients with rheumatoid arthritis (RA). This peptide is reactive with a	
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	The peptides were created by using standard solid phase synthesis, which	
CC	produced them as peptide amides. These sequences may be used in the	
CC	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as	
CC	for obtaining polyclonal and monoclonal antibodies	
SQ		
XX	Sequence 19 AA;	
Query Match	88.6%; Score 78; DB 2; Length 19;	
Best Local Similarity	89.5%; Pred.No.1.le-05;	
Matches 17; Conservative	0; Mismatches 2; Indels 0; Gaps 0	
OY	1 SHOESTXGRSGRSGRS 15 1 SHOESTXGRSGRXSGRS 19	
Dd		
RESULT 12		
AAM61509		
ID	AAM61509 standard; peptide; 19 AA.	
AC	AAM61509;	
DT	26-OCT-1998 (first entry)	
DE	Peptide cfc5, based on cDNA of a profilaggrin repeat.	
KM	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
KM	solid phase synthesis; peptide amide; polyclonal antibody;	
XX	monoclonal antibody.	
OS	Synthetic.	
OS	Homo sapiens.	
FH	Key Location/Qualifiers	
FT	Modified-site 16	
FT	/note="Citrulline"	
PX	MO9822503-A2.	
PD	28-MAY-1998.	
PF	14-NOV-1997; 97WO-NL000624.	
XX		
XX	15-NOV-1996; 96NL-01004539.	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
PI	Van Venrooij JWM, Schellekens GA, Raats JMH, Hoet RMA;	
XX	WPI; 1998-398613/34.	
DR		
PT	Peptide derived from an antigen recognised by autoantibodies - is	
FT		
PS		
XX		

[illegible]

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX
 SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 1 SHOESTXGRSGRSGSGS 19

RESULT 14
 AAM61510
 ID AAM61510 standard; peptide; 19 AA.

AC AAM61510;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 7 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"

XX MO9822503-A2;
 XX
 PD 28-MAY-1998.

XX PF 14-NOV-1997; 97MO-NL000624.
 XX PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
 DR MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 1.1e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 1 SHOESTXGRSGRSGSGS 19

RESULT 15
 AAM61520
 ID AAM61520 standard; peptide; 21 AA.

AC AAM61520;
 XX
 DT 26-OCT-1998 (first entry)
 XX

DE Peptide XI based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 9 /note= "Citnulline"
 FT Modified-site 9 /note= "Citnulline"

XX MO9822503-A2.
 XX
 PD 28-MAY-1998.

XX PF 14-NOV-1997; 97MO-NL000624.
 XX PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 XX (TEME-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WTW, Schellekens GA, Raats JMH, Hoet RMA;
 DR MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX Disclosure; Fig 1; 19pp; English.

XX Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

SQ Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 3.6e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGSGS 19
 DB 4 HOESTXGRSGRSGSGS 21

Search completed: September 28, 2004, 06:24:18
 Job time : 56.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)

54,782 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	9	US-09-308-150-12
2	84	95.5	19	9	US-09-308-150-1
3	84	95.5	19	9	US-09-308-150-3
4	84	95.5	19	9	US-09-308-150-7
5	84	95.5	19	9	US-09-308-150-11
6	84	95.5	19	9	US-09-308-150-13
7	84	95.5	19	9	US-09-308-150-14
8	78	88.6	19	9	US-09-308-150-2
9	78	88.6	19	9	US-09-308-150-4
10	78	88.6	19	9	US-09-308-150-5
11	78	88.6	19	9	US-09-308-150-6
12	78	88.6	19	9	US-09-308-150-8
13	78	88.6	19	9	US-09-308-150-9
14	75	85.2	21	9	US-09-308-150-10
15	75	85.2	22	9	US-09-747-029A-22

16	51	58.0	477	15	US-10-161-927-62	Sequence 62, Appl
17	49	55.7	179	12	US-10-425-114-53545	Sequence 53545, A
18	48	54.5	123	16	US-10-767-701-32436	Sequence 32436, A
19	47	53.4	164	16	US-10-437-963-112419	Sequence 112419, A
20	47	53.4	436	14	US-10-156-761-103033	Sequence 103033, A
21	46	52.3	243	16	US-10-437-963-130322	Sequence 130322, A
22	45	51.1	612	12	US-10-424-599-230819	Sequence 230819, A
23	44	50.0	133	16	US-10-437-963-136367	Sequence 136367, A
24	44	50.0	155	16	US-10-437-963-161956	Sequence 161956, A
25	44	50.0	653	16	US-10-437-963-198539	Sequence 198539, A
26	44	50.0	653	9	US-09-746-801A-13	Sequence 13, Appl
27	44	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
28	44	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
29	44	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
30	43.5	49.4	19723	15	US-10-084-846A-5	Sequence 5, Appl1
31	43	48.9	232	16	US-10-437-963-183394	Sequence 183394, A
32	43	48.9	536	16	US-10-437-963-149523	Sequence 149523, A
33	43	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
34	43	48.9	649	12	US-10-282-122A-76596	Sequence 76596, A
35	43	48.9	694	14	US-10-152-548-2	Sequence 2, Appl1
36	42	47.7	436	15	US-10-369-493-8178	Sequence 8178, Ap
37	42	47.7	506	16	US-10-437-963-185974	Sequence 185974, A
38	42	47.7	588	16	US-10-437-963-177280	Sequence 177280, A
39	41	46.6	24	9	US-09-971-543-12	Sequence 12, Appl
40	41	46.6	50	14	US-10-029-386-27692	Sequence 27692, A
41	41	46.6	103	12	US-10-424-599-257534	Sequence 257534, A
42	41	46.6	136	16	US-10-437-963-178685	Sequence 178685, A
43	41	46.6	178	12	US-10-425-114-89976	Sequence 69976, A
44	41	46.6	184	9	US-09-925-300-1619	Sequence 1619, Ap
45	41	46.6	184	12	US-10-296-115-763	Sequence 763, App

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Walthertus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoelt, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09/308,150
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-12
Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0;
Gy 1 SHQESTXGRSGRSGSGS 19

Db 1 SH0ESTXGRSGRSGSGS 19

```

RESULT 2
US-09-308-150-1
Sequence 1, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-1

```

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 6.4e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTXGRSGRSGSGS 19
 Db 1 SH0ESTXGRSGRSGSGS 19

```

RESULT 3
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19

```

```

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

```

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 94.7%; Pred. No. 6.4e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SH0ESTXGRSGRSGSGS 19
 Db 1 SH0ESTXGRSGRSGSGS 19

```

RESULT 4
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308.150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

```

Query Match 95.5%; Score 84; DB 9; Length 19;
 Best Local Similarity 100.0%; Pred. No. 6.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SH0ESTXGRSGRSGSGS 19
 Db 1 SH0ESTXGRSGRSGSGS 19

```

RESULT 5
US-09-308-150-11
Sequence 11, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

```

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-11
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
        |||||
Db      1 SHOESTGRSGRSGSGS 19
```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-13
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
        |||||
Db      1 SHOESTGRSGRSGSGS 19
```

```
RESULT 7
US-09-308-150-14
```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-14
```

```
Query Match          95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy      1 SHOESTXGRSGRSGSGS 19
        |||||
Db      1 SHOESTGRSGRSGSGS 19
```

```

RESULT 8
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2
```

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
|||
1 SHOESTXGRSGRSGRSGS 19

RESULT 9
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
|||
1 SHOESTXGRSGRSGRSGS 19

RESULT 10
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 5.1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
|||
1 SHOESTXGRSGRSGRSGS 19

RESULT 11
US-09-308-150-6

Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGRSGS 19
|||
1 SHOESTXGRSGRSGRSGS 19

RESULT 12
US-09-308-150-8

Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-8

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. NO. 5.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 13
US-09-308-150-9
Sequence 9, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-9

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. NO. 5.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGSGS 19
Db 1 SHOESTXGRSGRSGSGS 19

RESULT 14
US-09-308-150-10
Sequence 10, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from
OTHER INFORMATION: Known cDNA sequences of human profilaggrin
NAME/KEY: DISULFID
LOCATION: (3)..(16)
US-09-308-150-10

Query Match 85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. NO. 0.00016;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HOESTXGRSGRSGSGS 19
Db 4 HOESTXGRSGRSGSGS 21

RESULT 15
US-09-747-029A-22
Sequence 22, Application US/09747029A
Patent No. US20020143143A1
GENERAL INFORMATION:
APPLICANT: Union, Ann
APPLICANT: Moereels, Henri
APPLICANT: Meheus, Lydie
TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: RHEUMATOID ARTHRITIS
FILE REFERENCE: 11362.0031NPU50 INNS:031
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US/09/747,029A
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: EP 00870195.5
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: EP 99870280.7
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 22
TYPE: PRT

us-09-308-150-7.open.rapp

Page 6

Page 6

Page 6

Page 6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-7
Perfect score: 88
Sequence: 1 SHQSTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/CTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	51.1	518	4	US-09-252-991A-25967
2	44	50.0	395	4	US-09-252-991A-21702
3	44	50.0	483	4	US-09-252-991A-19015
4	44	50.0	653	4	US-09-513-057C-13
5	43	48.9	629	4	US-09-252-991A-22901
6	42	47.7	85	4	US-09-252-991A-29339
7	41.5	47.2	96	4	US-09-252-991A-19216
8	41	46.6	120	4	US-09-252-991A-17025
9	41	46.6	237	4	US-09-252-991A-17927
10	41	46.6	255	4	US-09-252-991A-19598
11	41	46.6	256	4	US-09-252-991A-26244
12	41	46.6	406	4	US-09-252-991A-19857
13	41	46.6	421	4	US-09-252-991A-33326
14	41	46.6	722	4	US-09-252-991A-19216
15	41	46.6	724	4	US-09-252-991A-19216
16	41	46.6	745	4	US-09-252-991A-19216
17	41	46.6	2237	1	US-08-523-849-36
18	40	45.5	123	4	US-09-252-991A-28146
19	40	45.5	213	4	US-09-252-991A-32093
20	40	45.5	261	4	US-09-252-991A-29394
21	40	45.5	429	4	US-09-252-991A-29394
22	40	45.5	429	4	US-10-109-854-5
23	40	45.5	429	4	US-09-252-991A-19026
24	40	45.5	468	4	US-09-252-991A-19026
25	40	45.5	484	4	US-09-252-991A-19026
26	40	45.5	508	2	US-08-818-024-3
27	40	45.5	722	4	US-08-817-832B-32

28	39.5	44.9	191	4	US-09-252-991A-18626	Sequence 18626, A
29	39	44.3	147	4	US-09-252-991A-22332	Sequence 22332, A
30	39	44.3	148	4	US-09-252-991A-25505	Sequence 25505, A
31	39	44.3	176	4	US-09-252-991A-29512	Sequence 29512, A
32	39	44.3	195	4	US-09-252-991A-30082	Sequence 30082, A
33	39	44.3	409	4	US-09-252-991A-14217	Sequence 14217, A
34	39	44.3	491	4	US-09-252-991A-7836	Sequence 7836, Ap
35	39	44.3	778	4	US-09-252-991A-18605	Sequence 18605, A
36	39	44.3	1293	4	US-09-170-968-252	Sequence 292, App
37	39	44.3	1253	4	US-09-364-425B-57	Sequence 57, Appl
38	38.5	43.8	257	4	US-09-252-991A-22204	Sequence 22204, A
39	38.5	43.8	388	4	US-09-252-991A-30849	Sequence 30849, A
40	38	43.2	26	4	US-09-297-981-22	Sequence 22, Appl
41	38	43.2	46	4	US-09-297-981-20	Sequence 20, Appl
42	38	43.2	223	1	US-07-667-276A-10	Sequence 10, Appl
43	38	43.2	279	4	US-09-252-991A-32606	Sequence 32606, A
44	38	43.2	299	4	US-09-252-991A-21789	Sequence 21789, A
45	38	43.2	369	4	US-09-252-991A-20245	Sequence 20245, A

ALIGNMENTS

RESULT 1
US-09-252-991A-25967
Sequence 25967, Application US/09252991A
Parent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25967
LENGTH: 518
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25967

Query Match 51.1% Score 45; DB 4; Length 518;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 STXGRSGRSGRG 18
DB 412 SGGGRSGRSGRG 425
RESULT 2
US-09-252-991A-21702
Sequence 21702, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21702
LENGTH: 395
TYPE: PRT

ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-21702

Query Match	50.0%	Score 44;	DB 4;	Length 395;
Best Local	Similarity	50.0%;	Pred. No. 14;	
Matches	9;	Conservative	2;	Mismatches 7;
				Indels

Qy	1	SHQESTXGRSXGRSG	18
		:	
Db	251	SHQGASEGRGDHESGRPG	268

RESULT 3
US-09-252-991A-19015
Sequence 19015, Application US/09252991A

Query Match	50.0%;	Score 44;	DB 4;	Length 483;
Best Local Similarity	47.1%;	Pred. No. 17;		
Matches	8;	Conservative	2;	Mismatches 7; Indels 0; Gaps 0;

```

RESULT 4
US-09-513-057C-13
; Sequence 13, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 653
; TYPE: PR1
; ORGANISM: Cardamine oligosperma
; US-09-513-057C-13

```

```

Query Match:          50.0%      Score 44; DB 4; Length 653;
Best Local Similarity 52.6%;    Pred. No. 23;
Matches   10; Conservative     1; Mismatches       8; Indels    0; Gaps    0.
QY      1 SHOESTXGSKXKSGRSGS 19
        | | | | | : | | | |
Db      573 SRQVSTASASAGREGISGS 591

```

RESULT 5
US-09-252-991A-22901
; Sequence 22901, Application US/09252991A

```

: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 22901
: LENGTH: 629
: TYPE: prt
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-22901

```

Query Match	48.9%;	Score 43;	DB 4;	Length 629;
Best Local Similarity	47.1%;	Pred. No. 32;		
Matches	8;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      3 QESTXGRSXGRSGRGS 19
          ::|||:||||
Db     272 RQSHRGRAGRAGRGS 288
```

RESULT 6
 US-09-252-991A-29339
 ; Sequence 29339, Application US/09252991A
 ; Patent No.6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; PRIORITY FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 29339
 ; LENGTH: 85
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-29339

```
Query Match      47.7%; Score 42; DB 4; Length 85;
Best Local Similarity 80.0%; Pred. No. 6;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Caps 0;
```

QY	8	GRSXGRSGRS	17
	:		
Dp	73	GRAAGRSGRS	82

```

RESULT 7
US-09-252-991A-19216
: Sequence: 19216, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERGININOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
:

```


PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19216
LENGTH: 96
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19216

Query Match 47.2%; Score 41.5; DB 4; Length 96;
Best Local Similarity 62.5%; Pred. No. 8.2;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 5 STXGR-SXGRSGSGS 19
Db 75 SATGRGAGGAGSGS 90

RESULT 8
US-09-252-991A-17025
Sequence 17025, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17025
LENGTH: 120
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17025

Query Match 46.6%; Score 41; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 SXGRSGSGS 19
Db 7 SAGRAGSGS 16

RESULT 9
US-09-252-991A-17927
Sequence 17927, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17927
LENGTH: 237
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17927

Query Match 46.6%; Score 41; DB 4; Length 237;
Best Local Similarity 80.0%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 GRGXGRSGRS 17
Db 24 GRGGRGREGRA 33

RESULT 10
US-09-252-991A-19598
Sequence 19598, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19598
LENGTH: 255
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19598

Query Match 46.6%; Score 41; DB 4; Length 255;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 8 GRGXGRSGRS 19
Db 64 GRDGRGAGRAGA 75

RESULT 11
US-09-252-991A-26244
Sequence 26244, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26244
LENGTH: 256
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26244

Query Match 46.6%; Score 41; DB 4; Length 256;
Best Local Similarity 47.1%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 HOESTXGRSGSGSG 18
Db 97 HRRRTAGRAGRRPRRG 113

RESULT 12
US-09-252-991A-17927

```

US-09-252-991A-19857
; Sequence 19857, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19857
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19857

Query Match          46.6%; Score 41; DB 4; Length 406;
Best Local Similarity 44.4%; Pred. No. 43;
Matches      8; Conservative      2; Mismatches      8; Indels      0; Gaps      0;

QY      2  HSESTXGRSXYGRSGRSGS 19
DB      127 HGSRRGRDRAGRGGRHGA 144

RESULT 13
US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326

Query Match          46.6%; Score 41; DB 4; Length 421;
Best Local Similarity 47.1%; Pred. No. 44;
Matches      8; Conservative      1; Mismatches      8; Indels      0; Gaps      0;

QY      2  HSESTXGRSXYGRSGRSG 18
DB      13  HARTGTGTGRRGRGRSG 29

RESULT 14
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31

```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match          46.6%; Score 41; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 77;
Matches      9; Conservative      0; Mismatches      6; Indels      0; Gaps      0;

QY      5  STXGRSXYGRSGRSGS 19
DB      562 SPSGHSQGRGRGASGS 576

RESULT 15
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-890-2

Query Match          46.6%; Score 41; DB 4; Length 724;
Best Local Similarity 60.0%; Pred. No. 77;
Matches      9; Conservative      0; Mismatches      6; Indels      0; Gaps      0;

QY      5  STXGRSXYGRSGRSGS 19
DB      564 SPSGHSQGRGRGASGS 578

```

Search completed: September 28, 2004, 06:26:45
Job time : 15.4479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88

Sequence: 1 SHOESTXGRSGXSGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	84	95.5	2248	2 A35938
2	78	88.6	416	2 A32947
3	69	78.4	591	2 A45135
4	52	59.1	506	1 W2K147
5	49	55.7	822	1 T51049
6	46.5	52.8	135	2 A46398
7	45	51.1	2420	2 A84652
8	44	50.0	638	2 I53169
9	44	50.0	800	2 T02852
10	44	50.0	1829	2 T35681
11	43	48.9	306	2 T21220
12	43	48.9	694	2 S71786
13	43	48.9	849	2 A96592
14	43	48.9	1015	2 T42013
15	43	48.9	1507	2 B47328
16	42	47.7	136	2 T35632
17	42	47.7	176	2 T35632
18	42	47.7	772	2 T27512
19	41	46.6	471	2 T33997
20	41	46.6	825	2 UC4163
21	41	46.6	836	2 G84727
22	41	46.6	311	1 JDUVLR
23	40.5	46.0	311	2 T45683
24	40	45.5	118	2 A72654
25	40	45.5	129	2 S14984
26	40	45.5	150	2 E71185
27	40	45.5	151	2 F75145
28	40	45.5	163	2 T23076
29	40	45.5	269	2 B38095

30	40	45.5	272	2 A38900	T-cell-specific tr
31	40	45.5	303	2 JH0401	T-cell-specific tr
32	40	45.5	312	2 A31846	130k paracrystall
33	40	45.5	373	2 T02976	probable DNA bindi
34	40	45.5	399	2 T46259	hypothetical prote
35	40	45.5	439	2 A38160	chdF protein - Esc
36	40	45.5	462	4 S33798	FUS/CHOP mutant fu
37	40	45.5	526	1 S33799	RNA-binding protei
38	40	45.5	827	2 T13468	DNA-directed DNA p
39	40	45.5	842	1 JDUVLR	repair/recombinati
40	40	45.5	854	2 S61164	hypothetical prote
41	40	45.5	1232	2 S40766	coucan gene protei
42	40	45.5	2176	2 T13806	hypothetical prote
43	39	44.3	110	2 B70601	hypothetical prote
44	39	44.3	259	2 C69122	hypothetical prote
45	39	44.3	313	2 A28444	flaggrin precursor

ALIGNMENTS

RESULT 1

A35938

Profilaggrin - human (fragments)

C.Species: Homo sapiens (man)

C.Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

C.Accession: A35938

R.Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.

Biochemistry 29, 9432-9440, 1990

A.Title: Organization, structure, and polymorphisms of the human profilaggrin gene.

A.Reference number: A35938; MUID:91064347; PMID:2248957

A.Accession: A35938

A.Status: preliminary; not compared with conceptual translation

A.Molecule type: DNA

A.Residues: 1-2248 <GN>

A.Cross-references: GB:U02929

A.Genetics:

A.Gene: GDB:FLG

A.Cross-references: GDB:119912; OMIM:135940

A.Map position: 1q21-1q21

C.Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

C.Keywords: EF hand; epidermis; polymorphism; tandem repeat

F:246-569/Region: flaggrin repeat

F:570-893/Region: flaggrin repeat

F:1074-1397/Region: flaggrin repeat

F:1573-1896/Region: flaggrin repeat

Query Match 95.5%; Score 84; DB 2; Length 2248;

Best Local Similarity 89.5%; Pred.No. 2.1e-05;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19

DB 227 SHOESTXGRSGXSGRSGS 245

RESULT 2

A32947

Flaggrin precursor - human (fragment)

C.Species: Homo sapiens (man)

C.Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999

C.Accession: A32947

R.McKinlay-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc

Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989

A.Title: Characterization of a cDNA clone encoding human flaggrin and localization of c

A.Reference number: A32947; MUID:89296901; PMID:2740331

A.Accession: A32947

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-416 <MCK>

A.Cross-references: GB:M4355; MUID:9182604; PIDN:AA52454.1; PID:9182605

A>Note: the authors translated the codon CAC for residue 186 as Gln, and AAT for residue

C.Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 88.6%; Score 78; DB 2; Length 416;
 Best Local Similarity 84.2%; Pred. No. 3.8e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 |||||
 DB 7 SHQSTXGRSGRSGSGS 25

RESULT 3

A45135
 A:Protein - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135
 R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nirunskisiri, W.; Dale, B.A.
 U. Biol. Chem. 267, 23772-23781, 1992
 A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717
 A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:L01089; GB:M90967; NID:G190408; PIDN:AAA60177.1; PID:G553621
 A:Note: sequence extracted from NCBI backbone (NCBI:118773)
 C:Genetics:
 A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 78.4%; Score 69; DB 2; Length 591;
 Best Local Similarity 78.9%; Pred. No. 0.0015;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 |||||
 DB 449 SHQSTXGRSGRSGSGS 467

RESULT 4

W2RL47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D35324
 R:Kiyono, T.; Adachi, A.; Ishibashi, M.
 Virology 177, 401-405, 1990
 A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
 A:Reference number: A35324; MUID:90281611; PMID:2162112
 A:Accession: D35324
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M32305; NID:G333062; PIDN:AAA6979.1; PID:G333067
 C:Superfamily: papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 59.1%; Score 52; DB 1; Length 506;
 Best Local Similarity 52.6%; Pred. No. 0.77;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
 |||||

DB 342 SRENGTRGRGRGRGAGS 360

RESULT 5

T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12F1.10
 C:Species: Neurospora crassa
 C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049
 R:Schulte, U.; Aigdn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: Z25286
 A:Accession: T51049
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12F1.10
 A:Experimental source: BAC clone B12F1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B12F1.10
 A:Map position: 6
 A:Introns: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;
 Best Local Similarity 52.9%; Pred. No. 3.8;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTXGRSGRSGSGS 19
 |||||
 DB 414 RESASGRTRGRGRGRT 430

RESULT 6

A46398
 RNA-binding protein 1 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 24-Sep-1999
 C:Accession: A46398
 R:Kim, Y.J.; Zuo, P.; Manley, J.L.; Baker, B.S.
 Genes Dev. 6, 2568-2579, 1992
 A:Title: The Drosophila RNA-binding protein RBP1 is localized to transcriptionally active
 A:Reference number: A46398; MUID:94040720; PMID:11340470
 A:Accession: A46398
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-135 <KIT>
 A:Cross-references: GB:L04929; NID:G158223; PIDN:AAA28850.1; PID:G158224
 A:Note: authors translated the codon GCT for residue 65 as Arg
 C:Genetics: rbp1
 A:Gene: FlyBase:Rbp1
 A:Cross-references: FlyBase:FBgn0010252
 A:Introns: 106/2
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
 F:12-74/Domain: ribonucleoprotein repeat homology <RRM3>

Query Match 52.8%; Score 46.5; DB 2; Length 135;
 Best Local Similarity 52.2%; Pred. No. 1.7;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 4 ESTXGRS-----GXSGRSGS 19
 |||||
 DB 80 EMSXGRSRRRRRGSGSGSGS 102

RESULT 7

A84652
 hypothetical protein At2g25730 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: A84652
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.; P
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

enes, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: A84652

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-2420 <STO>

A/Cross-references: GB:AEO02093; NID:g4874311; PIDN:AA031373.1; GSPDB:GN00139

C/Genetics:

A/Map position: 2

Query Match 51.1%; Score 45; DB 2; Length 2420;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 STXGRSRGXSGRSGS 19
DB 1353 STSGRSGYGGGSSNS 1367

RESULT 8
153169
cytokeratin 2 - human

C/Species: Homo sapiens (man)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999

C/Accession: 153169

R/Collin, C.; Ouhayoun, J.

Differentiation 51, 137-149, 1992

A>Title: Suprabasal marker proteins distinguishing keratinizing squamous epithelia: Cyto

A/Reference number: 153169; MUID:93114504; PMID:1282112

A/Accession: 153169

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-638 <RBS>

A/Cross-references: GB:M99063; NID:g181389; PIDN:AAA5746.1; PID:g181390

C/Superfamily: cytoskeletal keratin

Query Match 50.0%; Score 44; DB 2; Length 638;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 QESTGRSRGXSGRSGS 17
DB 7 KKSFSGRSGSGSGRS 21

RESULT 9
T02852

probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)

C/Species: Leishmania major

C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000

C/Accession: T02852; H81462

R/Myler, P.J.

submitted to the EMBL Data Library, May 1998

A>Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.

A/Reference number: Z14740

A/Accession: T02852

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-800 <MTL>

A/Cross-references: EMBL:AB001274; NID:g3264850; PID:g2266920

R/Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, F.; Lemley, C.; Magness, C.;

Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999

A>Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c

A/Reference number: A81453; MUID:99178987; PMID:10077609

A/Accession: H81462

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-800 <PVL>

A/Cross-references: GB:AE001274; NID:g3264850; PIDN:AA031373.1; PID:g2266920; GSPDB:GN00

A/Experimental source: strain WHOM/IL/81/Friedlin

C/Genetics:
A/Gene: L1439.4
A/Map position: 1

Query Match 50.0%; Score 44; DB 2; Length 800;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HQESTGRSRGXSGRSGS 19
DB 429 HRDGVRLSTTASGRSGS 446

RESULT 10
T35681
probable sensory histidine kinase - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C/Accession: T35681

R/Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1998

A/Reference number: Z21587

A/Accession: T35681

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1829 <HAR>

A/Cross-references: EMBL:AL031031; PIDN:CA019849.1; GSPDB:GN00070; SCOEDB:SC7C7.03

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOEDB:SC7C7.03

Query Match 50.0%; Score 44; DB 2; Length 1829;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSRGXSGRSGS 19
DB 21 SNQGRSRGXSGRSGS 35

RESULT 11
T21220

hypothetical protein F21H7.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C/Accession: T21220

R/Gardner, A.

submitted to the EMBL Data Library, March 1997

A/Reference number: Z19393

A/Accession: T21220

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-306 <MTL>

A/Cross-references: EMBL:Z93379; PIDN:CA07568.1; GSPDB:GN00023; CESP:F21H7.5

A/Experimental source: clone F21H7

C/Genetics:

A/Gene: CESP:F21H7.5

A/Map position: 5

A/Introns: 28/3; 146/3; 232/3

C/Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHQESTGRSRGXSGRSG 18
DB 48 SHRSKTKNRSGSGSGSG 65

RESULT 12
S71786

wingless receptor precursor dfz2 - fruit fly (*Drosophila melanogaster*)

C:Species: Drosophila melanogaster
 C>Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 C/Accession: S71786; S78444
 R/Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang, Y.; Macke, J.P.; Andrew, Nature 382, 223-230, 1996
 A/Title: A new member of the frizzled family from Drosophila functions as a wingless receptor
 A/Reference number: S71786; MUID:96353971; PMID:8717036
 A/Accession: S71786
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-694 <BHA>
 A/Cross-references: EMBL:U65589
 A/Note: mRNA was also sequenced
 R/Bhanot, P.; Wang, Y.; Nathans, J. submitted to the EMBL Data Library, July 1996
 A/Reference number: S78444
 A/Accession: S78444
 A/Molecule type: DNA
 A/Residues: 1-416, 'T', 418-694 <BHW>
 A/Cross-references: EMBL:U65589; NID:91518050; PIDN:AA047273.1; PID:91518051
 C/Genetics:
 A/Gene: dfz2
 A/Cross-references: FlyBase:FBgn0016797
 C/Superfamily: fruit fly frizzled protein
 C/Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 30;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHOESTYGRSGSGSGSGS 19
 DB 182 SYTAGSGSGSGSGSGS 200

RESULT 13
 A96592
 hypochloral protein P14C21.55 (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: A96592
 R/Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chitt, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anssen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: A96592
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-849 <STO>
 A/Cross-references: GB:AE005173; NID:911055759; PIDN:AA028231.1; GSPDB:GNC00141
 C/Genetics:
 A/Gene: P14C21.55
 A/Map position: 1

Query Match 48.9%; Score 43; DB 2; Length 849;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 5 STXGRSGSGSGSGS 19
 DB 736 SSANRSGSGSGSGS 750

RESULT 14
 T42013
 frequency clock protein - Creopus spinulosus

C:Species: Creopus spinulosus
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
 C/Accession: T42013
 R/Lewis, M.T.; Morgan, L.W.; Feldman, J.F.
 Mol. Gen. Genet. 253, 401-414, 1997
 A/Title: Analysis of frequency (frq) clock protein homologs: evidence for a helix-turn-helix
 A/Reference number: 222024; MUID:97188515; PMID:9037100
 A/Accession: T42013
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1015 <LEW>
 A/Cross-references: EMBL:U25850; NID:9852501; PID:9852502; PIDN:AAA68072.1
 C/Genetics:
 A/Gene: frq
 A/Introns: 100/1

Query Match 48.9%; Score 43; DB 2; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 44;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 SHOESTYGRSGSGSGSGS 19
 DB 583 NHRKQKTHSTGDSGSGN 601

RESULT 15
 B47328
 natural killer cell tumor-recognition protein - mouse
 N/Alternate names: cyclophilin-related NK-tumor recognition protein; natural killer-tumor
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 05-Nov-1999
 C/Accession: B47328; I77662
 R/Anderson, S.K.; Gallinger, S.; Roder, J.; Frey, J.; Young, H.A.; Ortaldo, J.R.
 Proc. Natl. Acad. Sci. U.S.A. 90, 542-546, 1993
 A/Title: A cyclophilin-related protein involved in the function of natural killer cells.
 A/Reference number: A47328; MUID:93133824; PMID:8421688
 A/Accession: B47328
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1507 <AND>
 A/Cross-references: GB:L04289; NID:9192866
 A/Note: authors translated the codon AGT for residue 972 as Arg
 R/Rintret, A.; Anderson, S.K.
 Mol. Immunol. 30, 1307-1313, 1993
 A/Title: IL-2 regulates the expression of the NK-TR gene via an alternate RNA splicing me
 A/Reference number: I57820; MUID:94019422; PMID:8413330
 A/Accession: I77662
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 235-237;263-294 <RIN>
 A/Cross-references: GB:565998; NID:9425701; PIDN:AA028500.1; PID:9425702
 C/Genetics:
 A/Gene: NK-TR
 C/Superfamily: natural killer cell tumor-recognition protein; cyclophilin homology
 C/Keywords: alternative splicing; lymphocyte
 F/60-230/Domain: cyclophilin homology <CYP>

Query Match 48.9%; Score 43; DB 2; Length 1507;
 Best Local Similarity 81.8%; Pred. No. 64;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 RSRGSGSGSGS 19
 DB 736 RSSGSGSGSGS 746

Search completed: September 28, 2004, 06:15:02
 Job time: 14.0625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-8
Perfect score: 88
Sequence: 1 SH0ESTXGRSGXSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	88.6	416	1	FILA_HUMAN
2	52	59.1	506	1	VE2_HPV47
3	46.5	52.8	135	1	RBE1_DROME
4	44	50.0	638	1	K220_HUMAN
5	43	48.9	694	1	FRQ2_DROME
6	43	48.9	1015	1	FRQ2_DROME
7	43	48.9	1453	1	NCROR_MOUSE
8	42	47.7	176	1	SSB_TREPA
9	42	47.7	772	1	MR11_CAEEL
10	41	46.6	825	1	SE5_FAT
11	41	46.6	843	1	DPO1_HPV47
12	41	46.6	983	1	T144_HUMAN
13	40.5	46.0	1115	1	TBC2_CHLRE
14	40	45.5	150	1	RL19_PYPAB
15	40	45.5	151	1	RL19_PYPAB
16	40	45.5	330	1	DUSB_HUMAN
17	40	45.5	419	1	TCF7_MOUSE
18	40	45.5	526	1	FUS_HUMAN
19	40	45.5	660	1	DDXY_HUMAN
20	40	45.5	730	1	DPO1_HPV47
21	40	45.5	842	1	DPO1_HPV47
22	40	45.5	854	1	XRS2_YEAST
23	40	45.5	1232	1	YCOO5_CAEEL
24	39	44.3	526	1	YABA_YEAST
25	39	44.3	617	1	DBP1_YEAST
26	39	44.3	713	1	DBP1_YEAST
27	39	44.3	713	1	DDX4_RAT
28	39	44.3	824	1	TGM1_RAT
29	39	44.3	834	1	YME2_YEAST
30	39	44.3	1337	1	DEXT_STRDO
31	38.5	43.8	195	1	GRP3_DAUCA
32	38	43.2	124	1	SMDI_DROME
33	38	43.2	315	1	SOL2_YEAST

34	38	43.2	427	1	MYCN_SERCA	P26014 serinus can
35	38	43.2	502	1	VE2_HPV25	P36787 human papil
36	38	43.2	511	1	C7C4_ARATH	O64635 arabidopsis
37	38	43.2	648	1	FXR1_XENLA	P51115 xenopus lae
38	38	43.2	651	1	SEC9_YEAST	P40357 saccharomyc
39	38	43.2	1383	1	PRAX_RAT	O63425 rattus norv
40	38	43.2	1372	1	ERY1_SACER	O03133 saccharopol
41	38	43.2	3491	1	ERY1_SACER	O03131 saccharopol
42	37.5	42.6	857	1	VZA_CMYAT	O40977 cucumbei
43	37.5	42.6	1054	1	IF2_STRAU	P55875 stigmateila
44	37	42.0	123	1	LSM4_CAEEL	O19952 caenorhabdi
45	37	42.0	168	1	DH1_MAIZE	P12950 zea mays (m

ALIGNMENTS

RESULT 1
FILA_HUMAN STANDARD; PRT; 416 AA.
AC P26014: 1
ID FILA_HUMAN 1
AC P26014: 1
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE FILA_HUMAN precursor (Fragment).
GN FILA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69296901; PubMed=2740331;
RA McIntirey-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cannizzaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.;
RT "Characterization of a cDNA clone encoding human Filaggrin and
RT localization of the gene to chromosome region 1q21."
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RN [2]
RP CITRULLINATION.
RX MEDLINE=96374388; PubMed=8780679;
RA Sershan T., Kan S., Ogawa H., Marabe M., Asaga H.;
RT "Preferential determination of keratin Ki and filaggrin during the
RT terminal differentiation of human epidermis."
RL Blochum. Biophys. Res. Commun. 225:712-719(1996).
CC -1- FUNCTION: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -1- PFM: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 34 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.
CC -1- PFM: Undergoes delamination of some arginine residues
CC (citrullination).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M24355; AAA52454.1; -
CC PIR: A32947; A32947.
CC GeneW: HGNC:3748; FIG.
CC MIM: 135940; -
CC GO: GO:0005862; C:intermediate filament; NAS.
CC GO: GO:0005188; F:structural molecule activity; NAS.
CC GO: GO:0007275; P:development; NAS.
CC InterPro: IPR003303; Filaggrin.

```

DR Pfam; PF03516; Flaggrin; 2.
DR PRINTS; PR00487; FLAGGRIN.
KM Phosphorylation; Citrullination; Developmental protein.
FT NON TER
SO SEQUENCE 416 AA; 44105 MW; DEFA3218BA04332 CRC64;

Query Match 88.6%; Score 78; DB 1; Length 416;
Best Local Similarity 84.2%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOE5TGRGRGSGRSGS 19
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 7 SHOE5TGRGRGSGRSGS 25

RESULT 2
VE2 HPV47 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9028161; PubMed=2162112;
RA Kiyono T., Adachi A., Ishibashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
RT type 47 inferred from its DNA sequence.";
RL Virology 177:401-405 (1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE EARLY RESPONSE ELEMENT (5'-ACCONNNNGGT-3') PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2/E3'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY SPERITALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -.
DR PIR; D35324; W2WL47.
DR HSSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR Pfam; PF00551; E2_C_1.
DR Pfam; PF00508; E2_N_1.
DR Pfam; PF000672; E2_C_1.
DR Pfam; PD000678; E2_N_1.
DR Early protein; Transcription regulation; Activator; DNA-binding;
DR Trans-acting factor; DNA replication; Repressor; Nuclear protein.
KW SEQUENCE 506 AA; 57478 MW; 92C37F4BF8725065 CRC64;
SQ

```

```

Db 342 SRENTGRGRGSGRSGS 360
| : : | | | | | | | | | | | | | | | | | | | | | |
| : : | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
RBP1 DROME STANDARD; PRT; 135 AA.
AC Q02427; Q26271; Q9YGM8.
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein 1.
DE RBP1 OR CG17136.
GN RBP1 OR CG17136.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX MEDLINE=94040720; PubMed=1340470;
RA Kim Y.-J., Zuo P., Manley J.L., Baker B.S.;
RT "The Drosophila RNA-binding protein RBP1 is localized to
RT transcriptionally active sites of chromosomes and shows a functional
RT similarity to human splicing factor ASF/SF2.";
RL Genes Dev. 6:2569-2579 (1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abitl J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Beus P.V., Berman B.P., Bhandari D., Bolintiner S.,
RA Borkova D., Botchan M.R., Bouck J., Brotherton P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertire S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splitter B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yen R.-F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE OF 14-53 FROM N.A.
RX MEDLINE=93109300; PubMed=8417324;
RA Kim Y.-J., Baker B.S.;
RT "Isolation of RRM-type RNA-binding protein genes and the analysis of

```



```

RT their relatedness by using a numerical approach."
RL Mol. Cell. Biol. 13:174-183(1993).
RN
RP FUNCTION.
RX MEDLINE=95393975; PubMed=7664738;
RA Heinrichs V., Baker B.S.;
RT "The Drosophila SR protein RBPl contributes to the regulation of
RT doublesex alternative splicing by recognizing RBPl RNA target
RT sequences."
RL EMBL J. 14:3967-4000(1995).
CC -1- FUNCTION: Contributes to the activation of female-specific DSX
CC splicing in vivo by recognizing the RBPl target sequences within
CC the purine-rich polypyrimidine tract of the female-specific 3'
CC splice site.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=RBPl-A;
CC IsoId=Q02427-1; Sequence=Displayed;
CC Name=2; Synonyms=RBPl-B;
CC IsoId=Q02427-2; Sequence=VSP_005817;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- DEVELOPMENTAL STAGE: Found at all developmental stages.
CC -1- PTM: Extensively phosphorylated on serine residues in the RS
CC domain (Probable).
CC -1- SIMILARITY: Belongs to the splicing factor SR family.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L04929; AAA28850.1;
DR EMBL: AE003688; AAF5455.1;
DR EMBL: S51691; AAB24622.1;
DR PIR: A46398; A46398.
DR PIR: A48110; A48110.
DR FlyBase: FBgn0010252; Rbpl.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PFO0076; rrm.1.
DR PROSITE: PS50102; RRM.1.
DR PROSITE: PS00030; RRM_RNP.1; FALSE_NEG.
KM Nuclear protein; RNA-binding; mRNA splicing; Alternative splicing;
KM Phosphorylation.
FT DOMAIN 11 84 RNA-BINDING (RRM).
FT DOMAIN 82 135 ARG/SER-RICH (RS DOMAIN).
FT VARSPIC 107 135 Missing (in isoform 2).
FT CONFLICT 14 14 /FTID=VSP_005817.
FT CONFLICT 65 65 Y -> F (IN REF. 3).
FT CONFLICT 65 65 R -> A (IN REF. 1).
SQ SEQUENCE 135 AA; 15446 MW; 896DCE902518D991 CRC64;
Query Match 52.8%; Score 46.5; DB 1; Length 135;
Best Local Similarity 52.2%; Pred. No. 0.53;
Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
QY 4 ESTXGRSR-----GXSGRSGS 19
Db 80 EMSGGRSDRRRRRGSGSSGSGS 102
RESULT 4
K220 HUMAN STANDARD; PRT; 638 AA.
AC Q01546;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Keratin, type II cytoskeletal 2 oral (Cytokeratin 2p) (K2p)

```

```

DE (CK 2p).
GN KRT2p.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=93114504; PubMed=1282112;
RA Collin C., Ouhayoun J.P., Grund C., Franke W.W.;
RT "Suprabasal marker proteins distinguishing keratinizing squamous
RT epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium
RT and epidermis are different."
RL Differentiation 51:137-148(1992).
CC -1- FUNCTION: Probably contributes to terminal cornification.
CC -1- SUBUNIT: Heterotrimer of two type I and two type II keratins.
CC -1- DEVELOPMENTAL STAGE: Synthesized during maturation of epidermal
CC keratinocytes.
CC -1- MISCELLANEOUS: There are two types of cytoskeletal and
CC microfibrillar keratin: I (acidic; 40-55 kDa) [K9 to K20] and II
CC (neutral to basic; 56-70 kDa) [K1 to K8].
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M99063; AAA35746.1;
DR GO: GO:0005882; C:intermediate filament; NAS.
DR GO: GO:0005198; F:structural molecule activity; NAS.
DR GO: GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
DR InterPro: IPR001664; IF.
DR InterPro: IPR002857; Keratin_I.
DR InterPro: IPR003054; Keratin_II.
DR Pfam: PFO0038; filament.1.
DR PRINTS: PRO1248; TYPE1KERATIN.
DR PRINTS: PRO1276; TYPE2KERATIN.
DR PROSITE: PS00226; IF.1.
KM Intermediate filament; Coiled coil; Keratin; Phosphorylation.
FT DOMAIN 1 182 HEAD.
FT DOMAIN 183 492 ROD.
FT DOMAIN 493 638 TAIL.
FT DOMAIN 183 218 COIL 1A.
FT DOMAIN 219 237 LINKER 1.
FT DOMAIN 238 329 COIL 1B.
FT DOMAIN 330 353 LINKER 12.
FT DOMAIN 354 492 COIL 2.
FT MOD RES 57 57 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 638 AA; 65871 MW; 95743A8B872076AF CRC64;
Query Match 50.0%; Score 44; DB 1; Length 638;
Best Local Similarity 60.0%; Pred. No. 7.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 QESTXGRSGRSGSGS 17
Db 7 KXSFGSRGSGSGS 21
RESULT 5
FRZ2_DROME STANDARD; PRT; 694 AA.
AC Q9VWX3; Q94916; Q9VWX2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Frizzled protein 2 precursor (Frizzled-2) (dfzz).
DE FZ2 OR CG9739/CG14083.

```

CC Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydriidae; Drosophilidae; Drosophila.
 CC NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND BINDING TO WINGLESS THROUGH FZ DOMAIN.
 RX MEDLINE=96353971; PubMed=8717036;
 RA Bhano P., Brink M., Samos C.H., Hsieh J.C., Wang Y., Macke J.P.,
 RA Andrew D., Nathans J., Nusse R.;
 RT "A new member of the frizzled family from Drosophila functions as a
 RT WINGLESS receptor.";
 RL Nature 382:225-230 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731137;
 RA Adams M.D., Celtniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Baer A., Baxendale U., Bayraktaroglu L., Beasley B.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Cantler A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies A.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferriz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Nattai B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris U., Mostafaei A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors
 CC are coupled to the beta-catenin canonical signaling pathway, which
 CC leads to the activation of dishevelled proteins. Inhibition of
 CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation
 CC of Wnt target genes. A second signaling pathway involving PKC and
 CC calcium fluxes has been seen for some family members, but it is
 CC not yet clear if it represents a distinct pathway or if it can be
 CC integrated in the canonical pathway, as PKC seems to be required
 CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem
 CC to involve interactions with G-proteins. Required to coordinate
 CC the cytoskeletons of epidermal cells to produce a parallel array
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- DEVELOPMENTAL STAGE: Expression starts at stage 6 in all cells
 CC between 15 and 70 per cent of egg length, including the
 CC invaginating cells of the ventral furrow. Stripe pattern is
 CC emerging by early stage 8. From stage 9 and continuing throughout
 CC embryogenesis, expression is seen in the developing CNS. At stage

CC 10, expressed in 15 stripes in the presumptive head and trunk
 CC regions, in the posterior midgut primordium, in a subset of cells
 CC of anterior midgut invagination and in the procephalic lobe. At
 CC stage 12, expression declines in epidermis and increases in the
 CC midgut and visceral mesoderm. At stage 17, only expressed in the
 CC CNS, hindgut and dorsal vessel.
 CC -1- DOMAIN: Lys-Thr-X-X-X-Tyr motif is involved in the activation of
 CC the Wnt/beta-catenin signaling pathway (By similarity).
 CC -1- DOMAIN: The Fz domain is involved in binding with Wnt ligands.
 CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor
 CC family.
 CC -1- SIMILARITY: Contains 1 frizzled (Fz) domain.
 CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC DR EMBL: U65589; AAC47273.1; -;
 CC DR EMBL: AE003518; AAF49185.2; -;
 CC DR PIR: S71786; S71786.
 CC DR Flybase: FBgn0016797; fz2.
 CC DR GO: GO:0016021; C: integral to membrane, NAS.
 CC DR GO: GO:0017147; F: wnt-protein binding, IDA.
 CC DR GO: GO:0006928; P: cell motility, IMP.
 CC DR GO: GO:0007163; P: establishment and/or maintenance of cell po. . .; ISS.
 CC DR GO: GO:0008585; P: female gonad development, IMP.
 CC DR GO: GO:0016055; P: wnt receptor signaling pathway, IDA.
 CC DR InterPro: IPR000539; Frizzled.
 CC DR InterPro: IPR000024; Fz domain.
 CC DR InterPro: IPR000832; GPCR secretin.
 CC DR Pfam: PF01534; Frizzled_1.
 CC DR Pfam: PF01392; Fz_1.
 CC DR PRINTS: PR00489; FRIZZLED.
 CC DR SMART: SM00063; FRI: 1.
 CC DR PROSITE: PS50038; Fz: 1.
 CC DR PROSITE: PS50261; G_PROTEIN_RECEP_F2_4; 1.
 CC KW Multigene family; Receptor; G-protein coupled receptor; Transmembrane;
 CC Developmental protein; Wnt signaling pathway; Glycoprotein; Signal.
 CC FT SIGNAL 1 22
 CC FT CHAIN 23 694
 CC FT DOMAIN 23 315
 CC FT TRANSMEM 316 336
 CC FT DOMAIN 337 352
 CC FT TRANSMEM 353 373
 CC FT DOMAIN 374 397
 CC FT TRANSMEM 398 418
 CC FT DOMAIN 419 439
 CC FT TRANSMEM 440 460
 CC FT DOMAIN 461 482
 CC FT TRANSMEM 483 503
 CC FT DOMAIN 504 534
 CC FT TRANSMEM 535 555
 CC FT DOMAIN 556 584
 CC FT TRANSMEM 585 605
 CC FT DOMAIN 606 694
 CC FT DOMAIN 187 225
 CC FT DOMAIN 59 180
 CC FT SITE 608 613
 CC FT SITE 692 694
 CC FT CAROXYD 78 78
 CC FT CAROXYD 288 288
 CC FT CONFLICT 55 55
 CC FT CONFLICT 417 417
 CC FT CONFLICT 417 417
 CC SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAB0956 CRC64;
 CC Query Match 48.9%; Score 43; DB 1; Length 694;
 CC Best Local Similarity 52.6%; Pred. No. 12;
 CC Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 182 SYTEAGSGSGSGSGSGS 200

RESULT 6
 FRQ_CRESP STANDARD; PRT; 1015 AA.

AC 000586;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Frequency clock protein.
 FRQ.

OS Creopus spinulosus (Chromocrea spinulosa);
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
 OX NCBI_TaxID=110619;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FGSC 4896;
 RX MEDLINE=97188315; PubMed=90377100;
 RA Lewis M.T., Morgan L.W., Feldman J.F.;
 RT "Analysis of frequency (frq) clock gene homologs: evidence for a
 helix-turn-helix transcription factor";
 RL Mol. Gen. Genet. 253:401-414(1997).
 CC -!- FUNCTION: Circadian clock component involved in the generation of
 biological rhythms, in particular in rhythm stability; period
 length, and temperature compensation. Behaves as a negative
 element in circadian transcriptional loop (by similarity).
 CC -!- SIMILARITY: BELONGS TO THE FRQ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U25850; AAA68072.1; -
 DR EIR; T42013; T42013.
 KW Biological rhythms; Transcription regulation; Nuclear protein.
 FT DOMAIN 240 245
 FT DOMAIN 356 368
 FT DOMAIN 443 451
 FT DOMAIN 584 588
 FT DOMAIN 892 913
 FT SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 583 NHRKQXTHSTGSDSGSGSN 601

RESULT 7
 NKR_MOUSE STANDARD; PRT; 1453 AA.

AC P30415;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NK-tumor recognition protein (Natural-killer cells cyclophilin-
 related protein) (NK-TR protein).
 GN NKTR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9313924; PubMed=8421688;
 RA Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
 RA Ortaido J.R.;
 RT "A cyclophilin-related protein involved in the function of natural
 killer cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
 CC [2]
 RP REVISIONS TO C-TERMINUS.
 RC STRAIN=BALB/c; TISSUE=Blood;
 RA Anderson S.K.;
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Component of a putative tumor-recognition complex.
 CC Involved in the function of NK cells.
 CC -!- SIMILARITY: Contains 1 cyclophilin-like P1ase domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L04289; AAA37500.2; ALT_INIT.
 DR HSSP; Q27450; 1A33.
 DR MCD; MGI:97346; Nkr.
 DR InterPro; IPR002130; CSA_P1ase.
 DR Pfam; PF00160; PRO_Isomerase; 1.
 DR PRINTS; PRO0153; CSA_P1ase.
 DR PROSITE; PS00170; CSA_P1ase_1; 1.
 DR PROSITE; PS50072; CSA_P1ase_2; 1.
 KW Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
 FT DOMAIN 1 176
 FT DOMAIN 222 241
 FT DOMAIN 422 459
 FT DOMAIN 964 1003
 FT DOMAIN 198 273
 FT DOMAIN 468 565
 FT DOMAIN 658 812
 FT DOMAIN 1303 1453
 FT SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;

Query Match 48.9%; Score 43; DB 1; Length 1453;
 Best Local Similarity 81.8%; Pred. No. 26;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 RSRGXSGRSGS 19
 DB 682 RSSGSGRSGS 692

RESULT 8
 SSB_TREPA STANDARD; PRT; 176 AA.

AC O83101;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
 GN SSB OR TP0062.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.W., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,

```

RA McDonald L., Arltach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete";
RL Science 281:375-388(1998).
CC -1- FUNCTION: This protein is essential for replication of the
CC chromosome. It is also involved in DNA recombination and repair
CC (by similarity).
CC -1- SIMILARITY: Contains 1 SSB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001191; AAC65057.1; -.
DR PIR; F71370; F71370.
DR HSSP; P02339; IEYG.
DR TIGR; TP0062; -.
DR InterPro; IPR008994; Nucleic_acid_DB.
DR InterPro; IPR000424; SSB_protein.
DR Pfam; PF00436; SSB; 1.
DR TIGRFAMs; TIGR00621; ssb; 1.
DR PROSITE; PS50935; SSB; 1.
DR DNA-binding; DNA repair; DNA replication; Complete proteome.
KW DOMAIN
FT SEQUENCE 176 AA; 19043 MW; 47C49D2F79927D2C CRC64;
SQ
Query Match 47.7%; Score 42; DB 1; Length 176;
Best Local Similarity 62.5%; Pred. No. 4;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 ESTXGRSRGXSGRSGS 19
Db 128 ESTAGRYGTSSDERS 143
RESULT 9
MR11_CAEEL STANDARD; PRT; 772 AA.
ID MR11_CAEEL
AC Q23255;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Double-strand break repair protein mre-11.
GN MRE-11 OR ZC302.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP MEDLINE=21135651; PubMed=11238374;
RX Chin G.M., Villeneuve A.M.;
RT "C. elegans mre-11 is required for meiotic recombination and DNA
RT repair but is dispensable for the meiotic G(2) DNA damage
RT checkpoint.";
RL Genes Dev. 15:522-534(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kelly P.F.;
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
CC Possesses single-strand endonuclease activity and double-strand-
CC specific 3'-5' exonuclease activity. Also involved in meiotic DSB
CC processing.
CC -1- FUNCTION: Required for meiotic crossing over and chiasma

```

```

CC formation. Pachytene morphology and homolog pairing are normal.
CC Vital in long term for maintenance of reproductive capacity of
CC subsequent generations.
CC -1- COPACITOR: Manganese (By similarity).
CC -1- SUBUNIT: Forms a complex with rad-50 (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the MRE11/RAD32 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73978; CAA8292.1; -.
DR PIR; T27512; T27512.
DR GenOnline; 208612; -.
DR WormPep; ZC302.1; CR06573.
DR InterPro; IPR003701; DNA_repair.
DR InterPro; IPR004643; M-peptidase.
DR InterPro; IPR007281; Mre11_DNA_bind.
DR Pfam; PF00149; Metallophos; 1.
DR Pfam; PF04152; Mre11_DNA_bind; 1.
DR TIGRFAMs; TIGR00583; mre11; 1.
DR DNA_repair; Hydrolase; Nuclease; Endonuclease; Exonuclease;
KW Nuclear protein; Manganese; Meiosis.
FT MUTAGEN 139 139 E->K: IN MRE-11-ME41; DEFECTIVE IN
FT MEIOTIC CHROMOSOME DEGRADATION.
SQ SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;
QY 1 SHQSTXGRSRGXSGRSGS 19
Db 697 SKQPTGRGRGRGARGAGAS 715
RESULT 10
SES_RAT STANDARD; PRT; 825 AA.
ID SES_RAT
AC Q63003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SES antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Brain;
RX MEDLINE=96015159; PubMed=8537300;
RX Suzuki E., Kojima N., Yoshimura K., Uyemura K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein SES in the nervous system.";
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: D37934; BAA07153.1; -
 DR PIR: JG4163; JC4163.
 KW DNA-binding; Nuclear protein; Antigen.
 SQ SEQUENCE 825 AA; 86831 MW; AF667FE2PD55BDF CRC64;
 Query Match 46.6%; Score 41; DB 1; Length 825;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 8 GRSRGSGRSGS 19
 DB 589 GRGRGGRGGRGS 600
 RESULT 11
 ID DPOL_HPBVR STANDARD; PRT; 843 AA.
 AC P03157;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
 GN Hepatitis B virus (subtype adr).
 OS Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
 OC NCBI_TaxID=106820;
 RX MEDLINE=83168919; PubMed=6300776;
 RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishio K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype adr and adv."
 RL Nucleic Acids Res. 11:1747-1757(1983).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA)(N).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-phosphomonoester.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: V00867; -1 NOT_ANNOTATED_CDS.
 DR PIR: A00704; JDIYVR.
 DR InterPro: IPR001462; DNAPOL_viral_C.
 DR InterPro: IPR000201; DNAPOL_viral_N.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00242; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; rvtc_2.
 DR Pfam: PF000814; DNAPOL_viral_C; 1.
 DR Transferrase; RNA-directed DNA polymerase; DNA-directed DNA polymerase; Hydrolase; Nuclease; Endonuclease; DNA replication; DNA-binding.
 KW SEQUENCE 843 AA; 94400 MW; A6B2D49083C4E8B CRC64;
 SQ
 Query Match 46.6%; Score 41; DB 1; Length 843;
 Best Local Similarity 52.9%; Pred. No. 31;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 3 QESTXGRSGRSGS 19
 DB 226 QPQGGSWARGSGRSGS 242
 RESULT 12
 ID Y144_HUMAN STANDARD; PRT; 993 AA.
 AC Q14157;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein KIAA0144.
 GN KIAA0144.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1."
 RL DNA Res. 2:167-174(1995).
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: D63478; BAA09765.1; -
 DR InterPro: IPR000449; UBA_domain.
 DR Pfam: PF00627; UBA; 1.
 DR SMART: SM00165; UBA; 1.
 DR PROSITE: PS50030; UBA; 1.
 KW Hypothetical protein.
 FT DOMAIN 49 89 UBA.
 SQ SEQUENCE 983 AA; 103930 MW; 50579CFA2931EBA68 CRC64;
 QY 8 GRSRGSGRSG 18
 DB 178 GRGRGSGRSG 188
 RESULT 13
 ID TBCC2_CHLRE STANDARD; PRT; 1115 AA.
 AC Q8VXE3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tbc2 translation factor, chloroplast precursor.
 GN TBCC2.
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 OC Chlamydomonadaceae; Chlamydomonas.
 OC NCBI_TaxID=3055;
 RX MEDLINE=137C / CC-125;
 RA Auchincloss A.H., Zerges W., Perron K., Girard-Bascou J., Rochaix J.-D.;
 RT "Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbc mRNA in Chlamydomonas reinhardtii."
 RL J. Cell Biol. 157:953-962(2002).
 CC -1- FUNCTION: Required for expression of the chloroplast encoded psbc mRNA, most likely for translation initiation. Interacts with the 5'UTR of psbc.
 CC -1- SUBUNIT: Part of a 400 kDa complex which is not stably associated

```

CC      with RNA.
CC      -1- SUBCELLULAR LOCATION: Chloroplast stroma.
CC      -1- SIMILARITY: IN THE CENTRAL SECTION, TO CRP1 OF ZEA MAYS.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; AJ27966; CAP20867.1; -.
KW      Chloroplast; Transalt peptide; Repeat.
FT      TRANSIT 1
FT      CHAIN 1115
FT      DOMAIN 483 1068
FT      REPEAT 483 521
FT      REPEAT 607 645
FT      REPEAT 685 723
FT      REPEAT 724 763
FT      REPEAT 764 803
FT      REPEAT 804 842
FT      REPEAT 843 880
FT      REPEAT 990 1029
FT      REPEAT 1030 1068
FT      DOMAIN 187 210
FT      DOMAIN 253 256
FT      DOMAIN 271 326
FT      DOMAIN 370 377
FT      DOMAIN 883 1115
SQ      SEQUENCE 1115 AA; 114823 MW; 2239799B91C5D8F7 CRC64;

Query Match
Best Local Similarity 46.0%; Score 40.5; DB 1; Length 1115;
Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Cy      5 STXGRSG--XSGRSGS 19
Db      170 STSGRAGMGSGGRNGS 187

RESULT 14
RL19_PYPHO STANDARD; PRT; 150 AA.
AC      OS9437;
DT      15-DEC-1998 (Rel. 37, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DE      50S ribosomal protein L19E.
GN      RPL19E OR PH1759.
OS      Pyrococcus horikoshii.
CC      Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC      Pyrococcus.
CC      NCBI_TaxID=53953;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=OT3;
RX      MEDLINE=98344137; PubMed=9679194;
RA      Kawarayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA      Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA      Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtoku Y.,
RA      Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kusuda N., Oguchi A.,
RA      Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA      Masuchi Y., Shizuya H., Kikuchi H.;
RT      "Complete sequence and gene organization of the genome of a hyper-
RT      thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL      DNA Res. 5:55-76(1998).
CC      -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its

```

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; AP000007; BAA30873.1; -.
DR      PIR; B71185; B71185.
DR      InterPro; IPR000196; Ribosomal_L19e.
DR      Pfam; PF01280; Ribosomal_L19e; 1.
DR      ProDom; PD004823; Ribosomal_L19e; 1.
DR      PROSITE; PS00526; RIBOSOMAL_L19E; 1.
KW      Ribosomal protein; Complete proteome.
SQ      SEQUENCE 150 AA; 17997 MW; 412223A854825B69 CRC64;

Query Match
Best Local Similarity 45.5%; Score 40; DB 1; Length 150;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Cy      2 HOESTXGRSGXSGRSG 18
Db      67 HEQKKGRHGRGSGRKG 83

RESULT 15
RL19_PYPAB STANDARD; PRT; 151 AA.
AC      Q9V1V3;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      50S ribosomal protein L19E.
GN      RPL19E OR PYRAB09230 OR PAB2134.
OS      Pyrococcus abyssi.
CC      Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC      Pyrococcus.
CC      NCBI_TaxID=29292;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=GS / Orsay;
RX      MEDLINE=2251545; PubMed=12622808;
RA      Cohen G.N., Barde V., Flament D., Galperin M., Helling R., Lecompte O.,
RA      Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA      Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT      "An integrated analysis of the genome of the hyperthermophilic
RT      archaeon Pyrococcus abyssi."
RL      Mol. Microbiol. 47:1495-1512(2003).
CC      -1- SIMILARITY: Belongs to the L19E family of ribosomal proteins.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; AJ248284; CAB49245.1; -.
DR      PIR; F75145; F75145.
DR      InterPro; IPR000196; Ribosomal_L19e.
DR      Pfam; PF01280; Ribosomal_L19e; 1.
DR      ProDom; PD004823; Ribosomal_L19e; 1.
DR      PROSITE; PS00526; RIBOSOMAL_L19E; 1.
KW      Ribosomal protein; Complete proteome.
SQ      SEQUENCE 151 AA; 18187 MW; 85FC80644B437A0 CRC64;

Query Match
Best Local Similarity 45.5%; Score 40; DB 1; Length 151;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Cy      2 HOESTXGRSGXSGRSG 18
Db      67 HEQKKGRHGRGSGRKG 83

```

Tue Sep 28 07:11:39 2004

us-09-308-150-8.open.rsp

Page 9

Search completed: September 28, 2004, 06:05:51
Job time : 7.22396 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-8
Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	84	95.5	797	4	Q16824 homo sapien
2	84	95.5	990	4	Q15206 homo sapien
3	84	95.5	1218	4	Q05331 homo sapien
4	78	88.6	798	4	Q9H4U3 homo sapien
5	78	88.6	1084	4	Q01212 homo sapien
6	69	78.4	591	4	Q01720 homo sapien
7	69	78.4	687	4	Q9H4U2 homo sapien
8	64	72.7	465	4	Q03838 homo sapien
9	63	71.6	322	4	Q07370 homo sapien
10	49	55.7	822	3	Q9P312 neurospora
11	46.5	52.3	144	5	Q8T9K6 dirosophila
12	46	52.3	1142	13	Q7S2V2 xenopus lae
13	45	51.1	2464	10	Q9SL90 arabidopsis
14	44	50.0	285	5	Q9M3V3 dirosophila
15	44	50.0	638	4	Q7Z795 homo sapien
16	44	50.0	652	5	Q77168 apis mellif

17	44	50.0	800	5	Q15845 leishmania
18	44	50.0	936	16	Q82H71 streptomyce
19	44	50.0	1288	5	Q9V9Y3 dirosophila
20	44	50.0	1829	16	Q86808 streptomyce
21	43	48.9	185	10	Q8LNN6 oryza sativ
22	43	48.9	306	5	Q45386 caenorhabd
23	43	48.9	448	4	Q8NER3 homo sapien
24	43	48.9	452	4	Q8NPV4 homo sapien
25	43	48.9	820	12	Q69375 mouse cytom
26	43	48.9	849	10	Q9C7M2 arabidopsis
27	43	48.9	849	10	Q949W6 arabidopsis
28	43	48.9	966	11	Q8CBP6 mus musculu
29	42.5	48.3	240	16	Q87YL4 pseudomonas
30	42	47.7	259	4	Q86VY2 mus musculu
31	42	47.7	322	16	Q8G4B5 bifidobacte
32	42	47.7	481	16	Q7UJPP9 rhodospirill
33	42	47.7	709	17	Q8T9P1 methanosarc
34	42	47.7	983	11	Q8B0S3 mus musculu
35	42	47.7	983	11	Q812D4 mus musculu
36	42	47.7	1014	11	Q8BIT6 mus musculu
37	42	47.7	1015	11	Q8BIW4 mus musculu
38	42	47.7	1035	11	Q8CIG7 mus musculu
39	42	47.7	1067	11	Q8CIG7 mus musculu
40	42	47.7	1084	13	Q7S1B7 mus musculu
41	42	47.7	1105	11	Q8K102 mus musculu
42	42	47.7	1107	11	Q80X50 mus musculu
43	42	47.7	1112	11	Q8B701 mus musculu
44	42	47.7	1229	3	Q8B2H1 neurospora
45	41.5	47.2	958	5	Q9VC54 dirosophila

ALIGNMENTS

RESULT 1
Q16824 PRELIMINARY; PRT; 797 AA.

AC Q16824; 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE=91064347; PubMed=2248957;
RX Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinhart P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL: M60502; AAA63248.1; -;
DR GO: GO:0005198; Functional molecule activity; IEA.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF03516; Filaggrin; 4.
DR PRINTS: PR00487; FILAGGRIN.
FT NON TER 1
SQ SEQUENCE 797 AA; 85176 MW; 60B6184763BDA86B CRC64;

Query Match 95.5%; Score 84; DB 4; Length 797;
Best Local Similarity 89.5%; Pred. No. 88-06; 2; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTXGRSGRSGSGS 19
Db 427 SHQESTXGRSGRSGSGS 445

RESULT 2

Q15206 PRELIMINARY; PRT; 990 AA.
 ID Q15206
 AC Q15206;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinhert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]";
 RL Biochemistry 29:9432-9440(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91255199; PubMed=2043621;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinhert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene";
 RL Biochemistry 30:5814-5814(1991).
 DR EMBL; M60494; AAA63244.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON TER 990
 SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;
 Query Match 95.5%; Score 84; DB 4; Length 990;
 Best Local Similarity 89.5%; Pred. No. 1e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 DB 227 SHQSTXGRSGRSGSGS 245
 RESULT 3
 ID Q05331 PRELIMINARY; PRT; 1218 AA.
 AC Q05331;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE FILAGGRIN (PROFILAGGRIN) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FORESKIN;
 RX MEDLINE=93109348; PubMed=8417356;
 RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
 RA Steinhert P.M.;
 RT "Profilaggrin is a major epidermal calcium-binding protein.";
 RL Mol. Cell. Biol. 13:613-625(1993).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS
 CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

CC PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 CC REPEATS.
 DR EMBL; M69943; AAA64487.1; -;
 DR PIR; A48118; A48118.
 DR HSSP; P02593; 1CDM.
 DR GO; GO:0005856; C:cycloskeleton; NAS.
 DR GO; GO:0005509; F:calcium ion binding; TAS.
 DR GO; GO:0030154; P:cell differentiation; NAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
 DR InterPro; IPR001751; CapB_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF03516; Filaggrin; 6.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PR00487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KW Phosphorylation; Polyprotein; Developmental protein; Calcium-binding;
 KW Polymorphism.
 FT CA_BIND 19 32 SITE I (BY SIMILARITY).
 FT CA_BIND 62 73 SITE II (BY SIMILARITY).
 FT NON TER 1218 1218
 SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5265B19C2 CRC64;
 Query Match 95.5%; Score 84; DB 4; Length 1218;
 Best Local Similarity 89.5%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 DB 449 SHQSTXGRSGRSGSGS 467
 RESULT 4
 ID Q9H4U3 PRELIMINARY; PRT; 798 AA.
 AC Q9H4U3;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE DJL4N1.1.2 (Profilaggrin 3' end) (Fragment).
 GN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL356504; CAC13171.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 4.
 DR PRINTS; PR00487; FILAGGRIN.
 FT NON TER 1
 SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;
 Query Match 88.6%; Score 78; DB 4; Length 798;
 Best Local Similarity 84.2%; Pred. No. 8.6e-05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSGRSGSGS 19
 DB 428 SHQSTXGRSGRSGSGS 446
 RESULT 5
 ID Q01212 PRELIMINARY; PRT; 1084 AA.

AC 001212; 003840;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Profilaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.O., McBride O.W., Idler W.W., Steinert P.M.;
 RT "Organization, structure, and polymorphisms of the human profilaggrin
 gene [published erratum appears in Biochemistry 1991 Jun
 11;30(23):5814]."
 RT 11;30(23):5814.
 RL Biochemistry 29:9432-9440(1990).
 DR EMBL; M60503; AAA63243.1; JOINED.
 DR EMBL; M60501; AAA63243.1; JOINED.
 DR GO; GO:0005882; C:intermediate filament; NAS.
 DR GO; GO:0005198; P:development; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF03516; Filaggrin; 6.
 DR PRINTS; PRO0487; FILAGGRIN.
 DR NON TER
 FT SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 88.6%; Score 78; DB 4; Length 1084;
 Best Local Similarity 84.2%; Pred. No. 0.00012;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOEYXGRSGXSGSGS 19
 DB 64 SHOEYXGRSGXSGSGS 82

RESULT 6
 Q01720 PRELIMINARY; PRT; 591 AA.
 AC 001720;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
 GN FIG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=93054736; PubMed=1429717;
 RA Presland R.B., Haydock P.V., Fleckman P., Ntunusukit W., Dale B.A.;
 RT "Characterization of the human epidermal profilaggrin gene. Genomic
 organization and identification of an S-100-like calcium binding
 domain at the amino terminus."
 RT J. Biol. Chem. 267:23772-23781(1992).
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
 TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
 34 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
 CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
 PROTEOLYTICALLY CLEAVED.
 CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
 WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
 REPEATS.
 DR EMBL; L01089; AAA60177.1; -;
 DR EMBL; L01090; AAA60176.1; -;

DR PIR; A45135; A45135.
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR MIM; 135940; -;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; efhand; 1.
 DR Pfam; PF03516; Filaggrin; 2.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 KW Polymorphism.
 FT PROPEP 1 293
 FT CHAIN 294 467
 FT PROPEP 468 474
 FT CHAIN 475 591
 FT CA BIND 19 32
 FT CA BIND 62 73
 FT NON TER 591 591
 FT SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
 Best Local Similarity 78.9%; Pred. No. 0.0022;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOEYXGRSGXSGSGS 19
 DB 449 SHOEYXGRSGXSGSGS 467

RESULT 7
 Q094U2 PRELIMINARY; PRT; 687 AA.
 AC 094U2;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE DJ14N1.1.1 (Profilaggrin 5, end) (Fragment).
 GN FIG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laid G.;
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AL36504; CAC13172.1; -;
 DR PIR; A48118; A48118.
 DR HSSP; P80511; 1E8A.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001751; CABP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR003303; Filaggrin.
 DR Pfam; PF00036; efhand; 1.
 DR Pfam; PF03516; Filaggrin; 3.
 DR Pfam; PF01023; S_100; 1.
 DR PRINTS; PRO0487; FILAGGRIN.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CABP; 1.
 DR NON TER 687 687
 FT SEQUENCE 687 AA; 76659 MW; 8000363FBBF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
 Best Local Similarity 78.9%; Pred. No. 0.0025;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGS 19
DB 449 SHOESTRGRSGRSGS 467

RESULT 8

ID Q03838 PRELIMINARY; PRT; 465 AA.

AC Q03838;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE FILAGGRIN (PROFILAGGRIN) (Fragment).

GN FLG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=91064347; PubMed=2248957;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 29:9432-9440(1990).

RN [2]

R2 REVISIONS.

R3 MEDLINE=91255199; PubMed=2043621;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene.";

RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISMS: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.

CC -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

CC EMBL; M60499; AAA63246.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 3.

DR PRINTS; PR00487; FILAGGRIN.

FT NON_TER 1

FT NON_TER 465

FT SEQUENCE 465 AA; 50280 MW; C883744CSE134097 CRC64;

SO

Query Match 72.7%; Score 64; DB 4; Length 465;
Best Local Similarity 73.7%; Pred. No. 0.012;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGS 19
DB 227 SHOESARGRSGRSGS 245

RESULT 9

ID Q75370 PRELIMINARY; PRT; 322 AA.

AC Q75370;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Epidermal filaggrin (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;

RA Gital-Neuhäuser E., Durieux J.J., Arnaud M., Dalbon P., Sebbag M., Vincent C., Simon M., Seneu T., Masson-Bessiere C., Jolivet-Reynaud C., Jolivet M., Serre G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated anti-filaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deamination of arginine residues.";

RL J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC23559.1; -

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.

DR PRINTS; PR00487; FILAGGRIN.

FT NON_TER 1

FT NON_TER 322

FT SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

SO

Query Match 71.6%; Score 63; DB 4; Length 322;
Best Local Similarity 72.2%; Pred. No. 0.012;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGS 18
DB 305 SHOESTRGRSGRSGS 322

RESULT 10

ID Q99312 PRELIMINARY; PRT; 822 AA.

AC Q99312;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Related to nucleolar phosphoprotein.

GN B12P1.10.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCBI_Taxid=5141;

RN [1]

RP SEQUENCE FROM N.A.

RA Schulte U., Aign V., Hehseisel U., Brandt P., Fartmann B., Holland R., Nyakatura G., Nemes H.W., Mannhaupt G.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA German Neurospora genome project;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL390091; CAB98213.1; -

DR PIR; T51049; T51049.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR005054; RNA_rec.mot.

DR InterPro; IPR005120; Smg-4_UZF3.

DR Pfam; PF00076; rrm; 1.

DR Pfam; PF03467; Smg4_UZF3; 1.

DR SMART; SMO0360; RRM; 1.

DR PROSITE; PS0102; RRM; 1.

DR PROSITE; PS00030; RRM_RNP_1; 1.

DR SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

SO

Query Match 55.7%; Score 49; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 8.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 QESTXGRSGRSGS 19
DB 414 RESAGSTRGRSGRSGT 430

RESULT 11

Q8T9K6 PRELIMINARY; PRT; 144 AA.
 ID Q8T9K6
 AC Q8T9K6
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE GM02602P (CG17136-PD).
 NCBI OR CG17136.
 CN RBP1 OR CG17136.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 (1)
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guanin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer J.G., Champagne M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaborel G.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch L., Besley E.M.,
 RA Bailew R.M., Baus B.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Eranzela C.C., Ferraz C., Ferreira S., Flischnmann W.,
 RA Foster C., Gabriellian A.S., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodde A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon S., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter C., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhao G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.",
 RL Science 287:12185-2195(2000).
 (3)
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Barzon J., An H., Baldwin D., Barzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Matvei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome.",
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (4)
 RP SEQUENCE FROM N.A.
 RA Maier S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminler J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bertman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotation of Drosophila melanogaster genome.",
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (5)
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 (6)
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069252; AL139397.1; -
 DR EMBL; AB003688; AAN13487.1; -
 DR FlyBase; PBgn010252; Rbp1.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 SO SEQUENCE 144 AA; 16013 MW; A5B6D782CED5A46 CRC64;
 Query Match 52.8%; Score 46.5; DB 5; Length 144;
 Best Local Similarity 52.8%; Pred. No. 3.4;
 Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;
 4 ESTXGRSR-----GXSGRSGS 19
 Db 80 EMSSGRSDRRRGRGSGSGSGS 102
 RESULT 12
 Q7SZV2 PRELIMINARY; PRT; 1142 AA.
 ID Q7SZV2
 AC Q7SZV2
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Period 1.
 GN PER1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxId=8335;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Constantine C.M., Sutton B., Taylor M., Green C.B.,
 RT "Xenopus laevis period 1.",
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY340103; AAQ18143.1; -
 SO SEQUENCE 1142 AA; 123519 MW; 15BECB698D9E0B6 CRC64;
 Query Match 52.3%; Score 46; DB 13; Length 1142;
 Best Local Similarity 58.8%; Pred. No. 38;
 Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 1 SHQESTXGRSRXSGRS 17
 Db 805 SKGESNSGRRRRSGSGS 821

RESULT 13

Q9SL90 PRELIMINARY; PRT; 2464 AA.

AC Q9SL90; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Hypothetical protein.

GN A19225730.

OS Arabidopsis thaliana (mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Roming C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006053; AAD3133.2; -.

DR PIR; A84652; A84652.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR005797; Cyto b6_N.

DR InterPro; IPR008941; TPR-like.

DR PROSITE; PS00192; CYTOCHROME_B_HEME_1.

KW Hypothetical protein.

SQ SEQUENCE 2464 AA; 274770 MW; 63E0942007645BSP CRC64;

Query Match 51.1%; Score 45; DB 10; Length 2464;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 STXGRSGXSGRSGS 19

Db 1353 STSGRSGXGCGSNS 1367

RESULT 14

Q9W3V3 PRELIMINARY; PRT; 285 AA.

AC Q9W3V3; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE CCL4435 protein.

GN CG14435.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.;

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abiri J.F., Agbayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck U., Brokstein P., Brotlier P.,

RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman J.U., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kentison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Mervinov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacled J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirska R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

DR EMBL; AB034339; AAF46210.1;

DR FLYbase; FBgn0029911; CG14435.

SQ SEQUENCE 285 AA; 29822 MW; 95BDE3388D4B78D3 CRC64;

Query Match 50.0%; Score 44; DB 5; Length 285;

Best Local Similarity 47.1%; Pred. No. 19;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 HOESTXGRSGXSGRSG 18

Db 38 HOHSNPGQNRGNNGGCG 54

RESULT 15

Q7Z795 PRELIMINARY; PRT; 638 AA.

AC Q7Z795; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Keratin 2p.

GN KRT2B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

OX NCBI_TaxID=9606;

[1]

RP SEQUENCE FROM N.A.

RA Rogers M.A., Langein L., Smith F.J.D., McLean W., Winter H.,

RA Beckmann I., Praetzel S., Schweizer J.;

RT "Characterisation of new members of the human Type II keratin gene

RT family.";

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ564103; CAD91891.1; -.

KW Keratin.

SQ SEQUENCE 638 AA; 65840 MW; 6A312B8260BE06A9 CRC64;

Query Match 50.0%; Score 44; DB 4; Length 638;

Best Local Similarity 60.0%; Pred. No. 45;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QESTXGRSRGRSGRS 17
::| | | | |
Db 7 KTSFSGRSGGRSGRS 21

Search completed: September 28, 2004, 06:12:47
Job time : 41.1771 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88

Sequence: 1 SHQESTXGRSGXGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	96.6	19	2	AAW61515
2	84	95.5	19	2	AAW61508
3	84	95.5	19	2	AAW61512
4	84	95.5	19	2	AAW61517
5	84	95.5	19	2	AAW61505
6	84	95.5	19	2	AAW61514
7	84	95.5	19	2	AAW61516
8	84	95.5	19	2	AAW61506
9	78	88.6	19	2	AAW61507
10	78	88.6	19	2	AAW61509
11	78	88.6	19	2	AAW61511
12	78	88.6	19	2	AAW61513
13	78	88.6	19	2	AAW61510
14	78	88.6	19	2	AAW61512
15	75	85.2	22	4	AAW61520
16	75	85.2	22	4	AAW61520
17	60	68.2	330	2	AAW61525
18	60	68.2	330	2	AAW61525
19	60	68.2	330	2	AAW61525
20	60	68.2	330	2	AAW61525
21	51	58.0	441	4	ABG19110
22	50	56.8	447	6	ABO07142
23	46.5	52.8	135	4	ABW6181
24	45	51.1	1711	4	AAW79819
25	45	51.1	1951	4	AAW78835

26	44	50.0	285	4	ABB69359	Abb69359	Drosophila
27	44	50.0	1199	4	ABB58274	Abb58274	Drosophila
28	43	48.9	73	4	AAU54919	Aau54919	Propionib
29	43	48.9	73	6	AAW51438	Aaw51438	Propionib
30	43	48.9	452	6	AAO16179	Aao16179	Human pro
31	43	48.9	569	4	AAW79339	Aaw79339	Human pro
32	43	48.9	569	4	AAW79340	Aaw79340	Human pro
33	43	48.9	694	2	AAW31267	Aaw31267	Drosophila
34	43	48.9	694	4	ABB71797	Abb71797	Drosophila
35	43	48.9	797	5	ABP73371	Abp73371	Candida a
36	43	48.9	982	2	AAW13320	Aaw13320	Murine Na
37	43	48.9	982	6	ABU05134	Abu05134	Murine Na
38	43	48.9	1482	7	ADE58671	Ade58671	Rat Prote
39	43	48.9	1482	7	ADE58675	Ade58675	Rat Prote
40	42	47.7	176	6	ABU48456	Abu48456	Protein e
41	42	47.7	772	4	AAW64573	Aaw64573	Human Mre
42	41.5	47.2	957	4	ABB62605	Abb62605	Drosophila
43	41.5	47.2	1938	6	ABP76679	Abp76679	Streptomy
44	41	46.6	77	5	ABP06528	Abp06528	Human ORF
45	41	46.6	184	3	AAW57041	Aaw57041	Human pro

ALIGNMENTS

RESULT 1	AAW61515	AAW61515 standard; peptide: 19 AA.
XX	AAW61515;	
AC	AAW61515;	
XX	26-OCT-1998 (first entry)	
DT	26-OCT-1998 (first entry)	
XX	Peptide cfa, based on cDNA of a profilaggrin repeat.	
DE	Peptide cfa, based on cDNA of a profilaggrin repeat.	
XX	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
KW	solid phase synthesis; peptide amide; polyclonal antibody;	
KW	monoclonal antibody.	
XX	Synthetic.	
OS	Homo sapiens.	
XX	WO9822503-A2.	
PN	28-MAY-1998.	
XX	14-NOV-1997; 97WO-NL000624.	
PF	15-NOV-1996; 96NL-01004539.	
XX	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
PA	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
XX	Van Venrooij MW, Schellekens GA, Raats JMH, Hoet RMA,	
PI	WPI, 1998-398613/34.	
XX	Peptide derived from an antigen recognised by autoantibodies - is	
PT	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
PT	used in diagnosis of the disease.	
XX	Disclosure; Page 6, 19PP; English.	
XX	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of	
CC	the profilaggrin antigen which is recognised by autoantibodies from	
CC	patients with rheumatoid arthritis (RA). This peptide is reactive with a	
CC	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
CC	The peptides were created by using standard solid phase synthesis, which	
CC	produced them as peptide amides. These sequences may be used in the	
CC	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as	
CC	for obtaining polyclonal and monoclonal antibodies	
XX	Sequence 19 AA;	
SQ		

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.5e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

RESULT 2

AAW61508
 ID AAW61508 standard; peptide; 19 AA.

XX AAW61508;

XX 26-OCT-1998 (first entry)

XX Peptide cfc4, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

XX

AAW61512
 ID AAW61512 standard; peptide; 19 AA.

XX AAW61512;

XX 26-OCT-1998 (first entry)

XX Peptide cfc6, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

XX

XX

XX

XX

XX

XX

XX

AAW61517
 ID AAW61517 standard; peptide; 19 AA.

XX AAW61517;

XX 26-OCT-1998 (first entry)

XX Peptide cfc0, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

XX solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX Synthetic.

XX Homo sapiens.

XX Key

XX Modified-site

XX MO9822503-A2.

XX 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

XX MPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be

XX used in diagnosis of the disease.

XX Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

XX the profilaggrin antigen which is recognised by autoantibodies from

XX patients with rheumatoid arthritis (RA). This peptide is reactive with a

XX RA patient's autoimmune antibodies which are reactive with profilaggrin.

XX The peptides were created by using standard solid phase synthesis, which

XX produced them as peptide amides. These sequences may be used in the

XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

XX for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

XX

XX

XX

XX

XX

XX

XX

RESULT 3

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

RESULT 4

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHOESTXGRSRGXSGRSGS 19
 DB 1 SHOESTXGRSRGXSGRSGS 19

RESULT 5

Peptide cfc0, based on cDNA of a profilaggrin repeat.

KM Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW Solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 OS Synthetic.
 OS Homo sapiens.
 XX MO9822503-A2.
 XX 28-MAY-1998.
 PD 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI, 1998-398613/34.
 DR WPI, 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 SQ
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 8e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGXSGRSGS 19
 DB 1 SHQESTXGRSRGXSGRSGS 19
 RESULT 5
 AAM61505
 ID AAM61505 standard; peptide; 19 AA.
 XX
 AC AAM61505;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc1, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT
 XX
 PN MO9822503-A2.
 XX 28-MAY-1998.
 PD

PF 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI, 1998-398613/34.
 DR WPI, 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 SQ
 CC Sequences AAM61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGXSGRSGS 19
 DB 1 SHQESTXGRSRGXSGRSGS 19
 RESULT 6
 AAM61514
 ID AAM61514 standard; peptide; 19 AA.
 XX
 AC AAM61514;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cf, based on cDNA of a profilaggrin repeat.
 DE
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruiline"
 FT
 XX
 PN MO9822503-A2.
 XX 28-MAY-1998.
 PD 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
 PI WPI, 1998-398613/34.
 DR WPI, 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from CC patients with rheumatoid arthritis (RA). This peptide is reactive with a CC RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which CC produced them as peptide amides. These sequences may be used in the CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as CC for obtaining polyclonal and monoclonal antibodies

SO Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQESTXGRSRGSGRSGS 19
1 SHQESTXGRSRGSGRSGS 19

Db 1 SHQESTXGRSRGSGRSGS 19

RESULT 7
AAW61516
ID AAW61516 standard; peptide; 19 AA.

XX
AC AAW61516;

XX
DT 26-OCT-1998 (first entry)

XX
DE Peptide cfe, based on CDNA of a profilaggrin repeat.

XX
KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.

XX
OS Synthetic.

OS Homo sapiens.

XX
FN W09822503-A2.

XX
FD 28-MAY-1998.

XX
PF 14-NOV-1997; 97WO-NL000624.

XX
PR 15-NOV-1996; 96NL-01004539.

XX
PA (SCHE-) STICHTING SCHEIKINDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX
PI Van Venrooij WJW, Scheilekens GA, Raats JMF, Hoet RMA;
DR WPI; 1998-398613/34.

XX
PT Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from CC patients with rheumatoid arthritis (RA). This peptide is reactive with a CC RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which CC produced them as peptide amides. These sequences may be used in the CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as CC for obtaining polyclonal and monoclonal antibodies

SO Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
Best Local Similarity 89.5%; Pred. No. 8e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQESTXGRSRGSGRSGS 19
1 SHQESTXGRSRGSGRSGS 19

Db 1 SHQESTXGRSRGSGRSGS 19

RESULT 8
ABW97605
ID ABW97605 standard; protein; 1467 AA.

XX
AC ABW97605;

XX
DT 27-JUN-2002 (first entry)

XX
DE Novel human protein SEQ ID NO: 873.

XX
KW Human; anti-naemic; vulnary; anti-inflammatory; immunomodulator;
KW anti-infectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.

XX
OS Homo sapiens.

XX
PN W0200222660-A2.

XX
PD 21-MAR-2002.

XX
PF 10-SEP-2001; 2001WO-US026015.

XX
PR 11-SEP-2000; 2000US-00659671.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX
DR WPI; 2002-292408/33.

XX
DR N-PSDB; ABW32791.

XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.

XX
PS Example 2; SEQ ID NO 873; 509pp; English.

XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat stroke
CC infertility, to screen for drugs, to treat inflammatory conditions e.g.
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention

XX
SO Sequence 1467 AA;

Query Match 95.5%; Score 84; DB 5; Length 1467;
Best Local Similarity 89.5%; Pred. No. 5.6e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQESTXGRSRGSGRSGS 19
1 SHQESTXGRSRGSGRSGS 19

Db 449 SHQESTXGRSRGSGRSGS 467

RESULT 9
AAW61506
ID AAW61506 standard; peptide; 19 AA.

XX
AC AAW61506;

XX

DT 26-OCT-1998 (first entry)
 XX Peptide cfc2, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT
 FT
 PN WO9822503-A2.
 XX 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSRGSGSGS 19
 DB 1 SHOESTYGRSRGSGSGS 19
 RESULT 10
 AAW61507
 ID AAW61507 standard; peptide; 19 AA.
 XX
 AC AAW61507;
 XX
 DT 26-OCT-1998 (first entry)
 XX Peptide cfc3, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Modified-site 11 /note= "Citruilline"
 FT
 FT
 PN WO9822503-A2.
 XX 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;
 QY 1 SHOESTYGRSRGSGSGS 19
 DB 1 SHOESTYGRSRGSGSGS 19
 RESULT 11
 AAW61509
 ID AAW61509 standard; peptide; 19 AA.
 XX
 AC AAW61509;
 XX
 DT 26-OCT-1998 (first entry)
 XX Peptide cfc5, based on cDNA of a profilaggrin repeat.
 DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Modified-site 16 /note= "Citruilline"
 FT
 FT
 PN WO9822503-A2.
 XX 28-MAY-1998.
 XX 14-NOV-1997; 97WO-NL000624.
 XX 15-NOV-1996; 96NL-01004539.
 XX

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PS used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 84.2%; Pred. No. 7.6e-06;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGSGSGS 19
 DB 1 SHQESTXGRSRGSGSGS 19
 RESULT 12
 AAW61511
 ID AAW61511 standard; peptide; 19 AA.
 XX
 AC AAW61511;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DS Peptide cfc7, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 7 /note= "Citruilline"
 FT Modified-site 11 /note= "Citruilline"
 FT
 FT
 FT
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.
 XX
 PS Disclosure; Page 6; 19pp; English.
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ
 Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQESTXGRSRGSGSGS 19
 DB 1 SHQESTXGRSRGSGSGS 19
 RESULT 13
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.
 XX
 AC AAW61513;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc9, based on cDNA of a profilaggrin repeat.
 XX
 KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 7 /note= "Citruilline"
 FT Modified-site 16 /note= "Citruilline"
 FT
 FT
 FT
 XX W09822503-A2.
 XX
 XX 28-MAY-1998.
 XX
 XX 14-NOV-1997; 97WO-NL000624.
 XX
 XX 15-NOV-1996; 96NL-01004539.
 XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX WPI; 1998-398613/34.
 DR
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PS used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGRSGSGS 19
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 14
 AAW61510
 ID AAW61510 standard; peptide; 19 AA.

AC AAW61510;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfc6, based on cDNA of a proflaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

XX
 PN W09822503-A2.
 XX
 PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX
 DR WPI; 1998-398613/34.

XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX
 PS Disclosure; Page 6; 19pp; English.

XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX

SQ Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 7.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHQSTXGRSRGRSGSGS 19
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

AC AAW61520;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Peptide XI based on cDNA of a proflaggrin repeat.
 XX
 XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 XX monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 9 /note= "Citruilline"
 FT Modified-site 9 /note= "Citruilline"

XX
 PN W09822503-A2.
 XX
 PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.
 XX
 PR 15-NOV-1996; 96NL-01004539.

XX
 PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX
 PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
 XX
 DR WPI; 1998-398613/34.

XX
 PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX
 PS Disclosure; Fig 1; 19pp; English.

XX
 CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX

SQ Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 2.6e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOESTXGRSRGRSGSGS 19
 DB 4 HOESTXGRSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:18
 Job time : 54.1302 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds

(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-8

Perfect score: 88

Sequence: 1 SHOESTXGRSGXSGRSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 segs, 321558718 residues 1349238

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	85	96.6	19 9 US-09-308-150-12	Sequence 12, Appl
2	84	95.5	19 9 US-09-308-150-1	Sequence 1, Appl
3	84	95.5	19 9 US-09-308-150-4	Sequence 4, Appl
4	84	95.5	19 9 US-09-308-150-8	Sequence 8, Appl
5	84	95.5	19 9 US-09-308-150-11	Sequence 11, Appl
6	84	95.5	19 9 US-09-308-150-13	Sequence 13, Appl
7	84	95.5	19 9 US-09-308-150-14	Sequence 14, Appl
8	78	88.6	19 9 US-09-308-150-2	Sequence 2, Appl
9	78	88.6	19 9 US-09-308-150-3	Sequence 3, Appl
10	78	88.6	19 9 US-09-308-150-5	Sequence 5, Appl
11	78	88.6	19 9 US-09-308-150-6	Sequence 6, Appl
12	78	88.6	19 9 US-09-308-150-7	Sequence 7, Appl
13	78	88.6	19 9 US-09-308-150-9	Sequence 9, Appl
14	75	85.2	21 9 US-09-308-150-10	Sequence 10, Appl
15	75	85.2	22 9 US-09-747-029A-22	Sequence 22, Appl

16	50	56.8	477	15	US-10-161-927-62	Sequence 62, Appl
17	47	53.4	436	15	US-10-369-493-8178	Sequence 8178, Ap
18	46	52.3	123	16	US-10-767-701-32436	Sequence 32436, A
19	45	51.1	78	16	US-10-437-963-203073	Sequence 203073,
20	45	51.1	105	16	US-10-437-963-196208	Sequence 196208,
21	44	50.0	133	16	US-10-437-963-136367	Sequence 136367,
22	44	50.0	1936	14	US-10-156-761-11212	Sequence 11212, A
23	44	50.0	1087	9	US-09-918-905-24	Sequence 24, Appl
24	44	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
25	43.5	49.4	92	16	US-10-437-963-118183	Sequence 118183,
26	43	48.9	103	12	US-10-424-559-257534	Sequence 257534,
27	43	48.9	185	16	US-10-437-963-137528	Sequence 137528,
28	43	48.9	452	16	US-10-406-765A-2385	Sequence 2385, Ap
29	43	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
30	43	48.9	694	14	US-10-152-548-2	Sequence 2, Appl
31	43	48.9	797	14	US-10-032-585-7208	Sequence 7208, Ap
32	43	48.9	889	16	US-10-437-963-181020	Sequence 181020,
33	42	47.7	121	12	US-10-424-559-203241	Sequence 203241,
34	42	47.7	156	16	US-10-437-963-200525	Sequence 200525,
35	42	47.7	176	12	US-10-287-963-76350	Sequence 76350, A
36	42	47.7	179	12	US-10-425-114-53545	Sequence 53545, A
37	42	47.7	506	16	US-10-437-963-185974	Sequence 185974,
38	42	47.7	588	16	US-10-437-963-177280	Sequence 177280,
39	42	47.7	618	16	US-10-767-701-45284	Sequence 45284, A
40	42	47.7	772	15	US-10-359-493-6474	Sequence 6474, Ap
41	41.5	47.2	136	16	US-10-767-701-40697	Sequence 40697, A
42	41.5	47.2	19723	15	US-10-084-846A-5	Sequence 5, Appl
43	41	46.6	50	14	US-10-029-386-27692	Sequence 27692, A
44	41	46.6	178	12	US-10-425-114-69976	Sequence 69976, A
45	41	46.6	184	9	US-09-925-300-1619	Sequence 1619, Ap

ALIGNMENTS

RESULT 1
US-09-308-150-12
; Sequence 12, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known CDNA sequences of human profilaggrin
; US-09-308-150-12

Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 3.9e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 SHOESTXGRSGXSGRSGS 19

Db 1 SH0ESTAGRSRSGSGSGS 19

RESULT 2

US-09-308-150-1
Sequence 1, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 5.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SH0ESTYGRSGXSGSGSGS 19
Db 1 SH0ESTYGRSGXSGSGSGS 19

RESULT 3

US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 5.6e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SH0ESTXRSRSGXSGSGSGS 19
Db 1 SH0ESTRGRSGXSGSGSGS 19

RESULT 4

US-09-308-150-8
Sequence 8, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SH0ESTYGRSGXSGSGSGS 19
Db 1 SH0ESTYGRSGXSGSGSGS 19

RESULT 5

US-09-308-150-11
Sequence 11, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 19

;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; PRIOR FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-11

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 6
US-09-308-150-13
;; Sequence 13, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-13

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 7
US-09-308-150-14

;; Sequence 14, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 14
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: Known CDNA sequences of human profilaggrin
US-09-308-150-14

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.6e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTXGRSGXSGRSGS 19
DB 1 SHOESTGRSGRSGRSGS 19

RESULT 8
US-09-308-150-2
;; Sequence 2, Application US/09308150
;; Patent No. US20020137092A1
;; GENERAL INFORMATION:
;; APPLICANT: Van Ventrcoof, Waltherus Jacobus Wilhelmus
;; APPLICANT: Schellekens, Gerardus Antonius
;; APPLICANT: Raats, Jozef Maria Hendrik
;; APPLICANT: Hoet, Rene Michael Antonius
;; APPLICANT: Stichting Scheikundig Onderzoek Nederland
;; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
;; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
;; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
;; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
;; CURRENT APPLICATION NUMBER: US/09/308,150
;; CURRENT FILING DATE: 1999-09-30
;; PRIOR APPLICATION NUMBER: PCT/NL97/00624
;; PRIOR FILING DATE: 1997-11-14
;; PRIOR APPLICATION NUMBER: NL 1004539
;; PRIOR FILING DATE: 1996-11-15
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 19
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:Derived from
;; OTHER INFORMATION: Known CDNA sequences of human profilaggrin
;; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTRGRSGXSGSGS 19

RESULT 9
US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTRGRSGXSGSGS 19

RESULT 10
US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTRGRSGXSGSGS 19

RESULT 11
US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
TITLE OF INVENTION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTXGRSGXSGSGS 19

RESULT 12
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTXGRSGXSGSGS 19

RESULT 13
US-09-308-150-9
; Sequence 9, Application US/09308,150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 4.5e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

Cy 1 SHOESTXGRSGXSGSGS 19
Db 1 SHOESTXGRSGXSGSGS 19

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308,150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoet, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT APPLICATION NUMBER: US/09/308,150
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

Query Match      85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 HOESTXGRSGRGRSGSGS 19
Db 4 HOESTXGRSGRGRSGSGS 21

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Koezeels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NUS00 INNS:031
; CURRENT APPLICATION NUMBER: US/09/747,029A
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT
```

/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic Peptides
/ NAME/KEY: MOD_RES
/ LOCATION: (10)..(10)
/ OTHER INFORMATION: Xaa represents Citrulline
US-09-747-029A-22

Query Match 85.2%; Score 75; DB 9; Length 22;
Best Local Similarity 88.9%; Pred. No. 0.00015;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 HQESTXGRSRGSGSGS 19
Db 5 HQESTXGRSRGSGSGS 22

Search completed: September 28, 2004, 07:28:52
Job time : 112.526 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-8
Perfect score: 88
Sequence: 1 SH0ESTXGRSGRSGSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/5CTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	52.3	491	4 US-09-489-039A-7836	Sequence 7836, Ap
2	44	50.0	546	4 US-09-252-991A-24620	Sequence 24620, A
3	42	47.7	395	4 US-09-252-991A-21702	Sequence 21702, A
4	41	46.6	237	4 US-09-252-991A-23304	Sequence 23304, A
5	41	46.6	255	4 US-09-252-991A-19598	Sequence 19598, A
6	41	46.6	421	4 US-09-252-991A-32326	Sequence 32326, A
7	41	46.6	843	4 US-08-591-502B-45	Sequence 45, Appl
8	40	45.5	213	4 US-09-252-991A-32093	Sequence 32093, A
9	40	45.5	465	4 US-09-252-991A-18919	Sequence 18919, A
10	40	45.5	480	4 US-09-252-991A-22879	Sequence 22879, A
11	40	45.5	486	2 US-08-821-355A-8	Sequence 8, Appl
12	40	45.5	486	2 US-09-003-687A-8	Sequence 8, Appl
13	40	45.5	486	3 US-09-136-605-8	Sequence 18, Appl
14	40	45.5	660	3 US-09-058-489-18	Sequence 91, Appl
15	40	45.5	660	3 US-09-058-489-91	Sequence 787, App
16	40	45.5	660	4 US-09-976-594-787	Sequence 49, Appl
17	40	45.5	730	4 US-08-591-502B-49	Sequence 50, Appl
18	40	45.5	842	4 US-08-591-502B-50	Sequence 51, Appl
19	40	45.5	842	4 US-08-591-502B-51	Sequence 51, Appl
20	40	45.5	843	4 US-09-119-528A-2	Sequence 48, Appl
21	40	45.5	843	4 US-08-591-502B-48	Sequence 59, Appl
22	40	45.5	843	4 US-08-591-502B-59	Sequence 46, Appl
23	40	45.5	845	4 US-08-591-502B-46	Sequence 292, App
24	40	45.5	1293	4 US-09-170-496D-292	Sequence 57, Appl
25	39.5	44.9	191	4 US-09-364-425B-57	Sequence 18626, A
26	39.5	44.9	191	4 US-09-252-991A-18626	Sequence 18626, A
27	39	44.3	250	4 US-09-252-991A-31343	Sequence 31343, A

28	39	44.3	476	4 US-09-252-991A-18968	Sequence 18968, A
29	39	44.3	550	4 US-09-252-991A-25062	Sequence 25062, A
30	38	43.2	116	4 US-09-489-039A-13128	Sequence 13128, A
31	38	43.2	146	4 US-09-252-991A-21166	Sequence 21166, A
32	38	43.2	175	4 US-09-252-991A-19711	Sequence 19711, A
33	38	43.2	215	4 US-09-252-991A-10077	Sequence 30077, A
34	38	43.2	329	4 US-09-252-991A-28988	Sequence 28088, A
35	38	43.2	339	4 US-09-252-991A-21715	Sequence 21715, A
36	38	43.2	369	4 US-09-252-991A-20245	Sequence 22599, A
37	38	43.2	430	4 US-09-252-991A-22599	Sequence 3, Appl
38	38	43.2	508	2 US-08-818-024-3	Sequence 3, Appl
39	38	43.2	508	3 US-09-334-775A-3	Sequence 25967, A
40	38	43.2	518	4 US-09-252-991A-25967	Sequence 10, Appl
41	38	43.2	580	3 US-08-482-677-10	Sequence 10, Appl
42	38	43.2	580	4 US-10-033-174-10	Sequence 18036, A
43	38	43.2	638	4 US-09-252-991A-18036	Sequence 18036, A
44	38	43.2	735	4 US-09-252-991A-18657	Sequence 18657, A
45	38	43.2	774	4 US-09-252-991A-16789	Sequence 16789, A

ALIGNMENTS

```

RESULT 1
US-09-489-039A-7836
Sequence 7836, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,20040001
CURRENT FILING DATE: US/09/489,039A
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7836
LENGTH: 491
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7836

Query Match      52.3%, Score 46, DB 4, Length 491,
Best Local Similarity 69.2%, Pred. No. 10;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 STXGRSGRSGRS 17
      : |||||
DB      217 TASGRSGRSGRS 229

RESULT 2
US-09-252-991A-24620
Sequence 24620, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT FILING DATE: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24620
LENGTH: 546
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24620

```

ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor


```

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: NO PCT/US94/08665
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauber
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45

Query Match          46.6%; Score 41; DB 4; Length 843;
Best Local Similarity 52.9%; Pred. No. 1,1e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 QESTXGRSRGXSGRSGS 19
DB 226 QPQGGMARKSGRSGS 242

RESULT 8
US-09-252-991A-32093
; Sequence 32093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32093
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32093

Query Match          45.5%; Score 40; DB 4; Length 213;
Best Local Similarity 47.1%; Pred. No. 40;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```

QY 3 QESTXGRSRGXSGRSGS 19
DB 182 RETTPGRPGDGGRGRT 198

RESULT 9
US-09-252-991A-18919
; Sequence 18919, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18919
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18919

Query Match          45.5%; Score 40; DB 4; Length 465;
Best Local Similarity 47.4%; Pred. No. 87;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGXSGRSGS 19
DB 83 SCQTGCRGALRGSGRAGS 101

RESULT 10
US-09-252-991A-22879
; Sequence 22879, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22879
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22879

Query Match          45.5%; Score 40; DB 4; Length 480;
Best Local Similarity 46.7%; Pred. No. 90;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGXSG 15
DB 52 SKXSSAGRGARSTG 66

RESULT 11
US-08-821-355A-8
; Sequence 8, Application US/08821355A
; Patent No. 5851775
; GENERAL INFORMATION:
```

APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,355A
FILING DATE: 20-MAR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107,05064
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5851775e
US-08-821-355A-8
Query Match 45.5%; Score 40; DB 2; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 HOESTXGRSGXSG 15
DB 383 HOESTGCKRNAFG 396
RESULT 12
US-09-003-687A-8
Sequence 8, Application US/09003687A
Patent No. 5998600
GENERAL INFORMATION:
APPLICANT: Barker, Nick
APPLICANT: Clevers, Hans
APPLICANT: Korinek, Vladimir
APPLICANT: Morin, Patrice
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
APPLICANT: Sparks, Andrew
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,687A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/821,355
FILING DATE: 20-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: 1107,05064
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 97430 BMB UT
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998600e
US-09-003-687A-8
Query Match 45.5%; Score 40; DB 2; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 HOESTXGRSGXSG 15
DB 383 HOESTGCKRNAFG 396
RESULT 13
US-09-136-605-8
Sequence 8, Application US/09136605A
Patent No. 6140052
GENERAL INFORMATION:
APPLICANT: He, Tong-Chuan
APPLICANT: Kinzler, Kenneth
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: Beta Catenin, TCF-4, and APC Interact to
FILE REFERENCE: 1107,75741
CURRENT APPLICATION NUMBER: US/09/136,605A
CURRENT FILING DATE: 1998-08-20
EARLIER APPLICATION NUMBER: 08/821,355
EARLIER FILING DATE: 1997-03-20
EARLIER APPLICATION NUMBER: 09/003,687
EARLIER FILING DATE: 1998-01-06
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 486
TYPE: PRT
ORGANISM: Homo sapiens
US-09-136-605-8
Query Match 45.5%; Score 40; DB 3; Length 486;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 HOESTXGRSGXSG 15

Db 383 H0ESTTGKXNARC 396

```

RESULT 14
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lam, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-18

```

```

Query Match          45.5%; Score 40; DB 3; Length 660;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 5 STXGRSGXSGRSG 18
Db 609 SGFGASRGSSSRSG 622

```

RESULT 15
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; EARLIER FILING DATE: 1998-04-10
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-91

```

```

Query Match          45.5%; Score 40; DB 3; Length 660;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 5 STXGRSGXSGRSG 18
Db 609 SGFGASRGSSSRSG 622

Search completed: September 28, 2004, 06:26:45
Job time : 14.4479 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 13.0625 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-9

Perfect score: 88

Sequence: 1 SHQESTXGRSRGRSGXSGS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

1: PIR_78:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	90.9	2248	2	A35938
2	78	88.6	416	2	A32947
3	69	78.4	591	2	A45135
4	52	59.1	506	1	W2W47
5	49	55.7	822	2	T51049
6	44	50.0	745	2	G01025
7	44	50.0	774	2	I48609
8	44	50.0	1119	2	T16720
9	43	48.9	306	2	T21220
10	43	48.9	694	2	S71786
11	43	48.9	843	2	H86209
12	43	48.9	849	2	A96592
13	43	48.9	1015	2	I42013
14	42	47.7	136	2	T35632
15	42	47.7	176	2	F71370
16	42	47.7	308	2	A29379
17	42	47.7	315	2	S53889
18	42	47.7	772	2	T27512
19	42	47.7	2233	2	S63347
20	42	47.7	2420	2	A84652
21	41	46.6	167	2	T52599
22	41	46.6	174	2	T52600
23	41	46.6	373	2	T02976
24	41	46.6	471	2	T33997
25	41	46.6	528	2	G02127
26	41	46.6	759	2	P86362
27	41	46.6	825	2	UC4163
28	41	46.6	836	2	G84727
29	41	46.6	880	2	T04523

30	40.5	46.0	625	2	A34615	profilaggrin - rat
31	40	45.5	259	2	C69122	hypothetical prote
32	40	45.5	312	2	A31846	130K paracrystall
33	40	45.5	331	2	F63142	F631.5 protein - A
34	40	45.5	346	2	A82623	Tota protein Xf189
35	40	45.5	399	2	T46259	hypothetical prote
36	40	45.5	462	4	S33798	FMS/CHOP mutant fu
37	40	45.5	467	2	T01462	hypothetical prote
38	40	45.5	526	1	S33799	RNA-binding protei
39	40	45.5	526	2	A41698	cell division cont
40	40	45.5	807	2	T02738	probable AAA-type
41	40	45.5	854	2	S61164	repeat/recombinati
42	40	45.5	1557	2	T02859	probable serine/th
43	38.5	44.9	135	2	A46398	RNA-binding protei
44	38.5	44.9	286	2	S50855	neurotrophin-6 - s
45	39.5	44.9	1895	2	T06609	disease resistance

ALIGNMENTS

RESULT 1

A35938

profilaggrin - human (fragments)

C:Species: Homo sapiens (man)

C>Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999

C:Accession: A35938

R:Gan, S.Q.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.

Biochemistry 29, 9432-9440, 1990

A>Title: Organization, structure, and polymorphisms of the human profilaggrin gene.

A:Reference number: A35938; PMID:2248957

A:Accession: A35938

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2248 <GNN>

A:Cross-references: GB:U02929

C:Genetics:

A:Gene: GDB:FLG

A:Cross-references: GDB:119912; OMIM:135940

A:Map position: 1q21-q21

C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

F:246-569/Region: profilaggrin repeat

F:570-893/Region: profilaggrin repeat

F:1074-1397/Region: profilaggrin repeat

F:1573-1896/Region: profilaggrin repeat

Query Match 90.9%; Score 80; DB 2; Length 2248; Best local similarity 84.2%; Pred. No. 0.00011; Matches 16; Conservative 1; Mismatches 2; Gaps 0;

Qy 1 SHQESTXGRSRGRSGXSGS 19
Db 551 SHQESTXGRSRGRSGXSGS 569

RESULT 2

A32947

profilaggrin precursor - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 20-Dec-1989 #sequence_revision 04-Sep-1992 #text_change 29-Sep-1999

C:Accession: A32947

R:McKinley-Grant, L.V.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cannizzaro, L.; Croc

Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989

A>Title: Characterization of a cDNA clone encoding human profilaggrin and localization of th

A:Reference number: A32947; PMID:89296901; PMID:2740331

A:Accession: A32947

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <MCK>

A:Cross-references: GB:M24355; NID:G182604; PIDN:AAA52454.1; PID:G182605

A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue

C:Genetics:

A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 88.6%; Score 78; DB 2; Length 416;
Best Local Similarity 84.2%; Pred. No. 4.7e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
DB 7 SHOESTGRSGRSGSGS 25

RESULT 3

A45135
A:Protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C/Accession: A45135
R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nirunskiri, W.; Dale, B.A.
J. Biol. Chem. 267, 23772-23781, 1992

A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
A:Reference number: A45135; MUID:93054736; PMID:1429717
A:Accession: A45135
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-591 <PRE>
A:Cross-references: GB:L01089; GB:M90697; NID:9190408; PID:AAA60177.1; PID:9553621
A:Note: sequence extracted from NCBI backbone (NCBIP:118773)
C:Genetics:

A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21

C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 78.4%; Score 69; DB 2; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.0019;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
DB 449 SHOESTGRSGRSGSGS 467

RESULT 4

M2ML47
E2 protein - human papillomavirus type 47
C:Species: human papillomavirus type 47
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C/Accession: D53324
R:Kiyono, T.; Adachi, A.; Ishibashi, M.
Virology 177, 401-405, 1990

A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infe
A:Reference number: A35324; MUID:90281611; PMID:2162112
A:Accession: D53324

A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-506 <KIT>
A:Cross-references: GB:M2305; NID:9333062; PID:AAA46979.1; PID:9333067
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.89;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19

DB 342 SREGTRGRGRGRGRGS 360

RESULT 5

T51049
related to nucleolar phosphoprotein [imported] - Neurospora crassa
N/Alternate names: protein B12FL10
C:Species: Neurospora crassa
C/Accession: T51049
C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
R:Schulte, U.; Aign, V.; Hehse, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51049

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-822 <SCH>
A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12FL10
A:Experimental source: BAC clone B12FL; strain OR74A
C:Genetics:
A:Gene: NCSP:B12FL10
A:Map position: 6
A:Introns: 80/2

Query Match 55.7%; Score 49; DB 2; Length 822;
Best Local Similarity 52.9%; Pred. No. 4.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QESTYGRSGRSGSGS 19
DB 414 RESAGRTGRGRGRGT 430

RESULT 6

G01025
serine/threonine protein kinase - human
C:Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998
C/Accession: G01025
R:Navarro, E.
submitted to the EMBL Data Library, April 1996

A:Reference number: H00564
A:Accession: G01025
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-745 <NAV>
A:Cross-references: EMBL:X97630; NID:91310674
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F:18-271/Domain: protein kinase homology <KIN>

Query Match 50.0%; Score 44; DB 2; Length 745;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGSGS 19
DB 585 SPSGSGRGRGRGS 599

RESULT 7

I48609
probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
C:Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
C/Accession: I48609; S91333
R:Ingilis, J.D.; Lee, M.; Hill, R.E.
Mamm. Genome 4, 401-403, 1993

A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
A:Reference number: I48609; MUID:93364122; PMID:8358177
A:Accession: I48609
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-774 <INGI>

A:Cross-references: EMBL:X70764; NID:g57919; PIDN:CAA50040.1; PID:g57920
 R:Ingalls, J.D.; Lee, M.; Hill, R.E.
 Submitted to the EMBL Data Library, January 1993
 A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome
 A:Reference number: S3133
 A:Accession: S3133
 A:Molecule type: mRNA
 A:Residues: 1-698, 'X', '700-702', 'GRLPFAAKASCTSGTCYCAAMHQATRR', '731-774 <IMG2>
 A:Cross-references: EMBL:X70764
 C:Genetics:
 A:Gene: emk
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 C:51-304/Domain: protein kinase homology <XIN>
 F:59-67/Region: protein kinase ATP-binding motif

Query Match 50.0%; Score 44; DB 2; Length 774;
 Best Local Similarity 60.0%; Pred. No. 26;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 5 STXGRSGRSGXSGS 19
 Db 614 SPGCHGGRGASGS 628

RESULT 8
 T16720
 hypothetical protein R11G1.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16720
 R:Miller, N.
 Submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid R11G1.
 A:Reference number: Z18566
 A:Accession: T16720
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1119 <MIL>
 A:Cross-references: EMBL:U41016; NID:g1086652; PID:g1086656; PIDN:AAA82321.1; CESP:R11G1
 C:Genetics:
 A:Gene: CESP:R11G1.1
 A:introns: 24/1; 65/1; 107/1; 145/1; 500/1; 661/1; 693/3; 746/1; 802/2; 857/3; 926/3; 97

Query Match 50.0%; Score 44; DB 2; Length 1119;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 4 ESTXGRSGRSGXSGS 19
 Db 1008 EETTKENRREGGGS 1023

RESULT 9
 T21220
 hypothetical protein F21H7.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T21220
 R:Gardner, A.
 Submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19393
 A:Accession: T21220
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-306 <WIL>
 A:Cross-references: EMBL:Z93379; PIDN:CAB07586.1; GSPDB:GN00023; CESP:F21H7.5
 A:Experimental source: clone F21H7
 C:Genetics:
 A:Gene: CESP:F21H7.5
 A:Map position: 5
 A:introns: 28/3; 146/3; 232/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein F36D3.4

Query Match 48.9%; Score 43; DB 2; Length 306;
 Best Local Similarity 44.4%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGXSG 18
 Db 48 SHRSKTKRKSGKSGSG 65

RESULT 10
 S71786
 wingless receptor precursor dfz2 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Mar-2001
 C:Accession: S71786; S78444
 R:Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang, Y.; Macke, J.P.; Andrew,
 Nature 382, 225-230, 1996
 A:Title: A new member of the frizzled family from Drosophila functions as a wingless rec
 A:Reference number: S71786; MUID:96353971; PMID:8717036
 A:Accession: S71786
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-694 <BHA>
 A:Cross-references: EMBL:U65589
 A:Note: mRNA was also sequenced
 R:Bhanot, P.; Wang, Y.; Nathans, J.
 Submitted to the EMBL Data Library, July 1996
 A:Reference number: S78444
 A:Accession: S78444
 A:Molecule type: DNA
 A:Residues: 1-416, 'T', '418-694 <BHW>
 A:Cross-references: EMBL:U65589; NID:g1518050; PIDN:AA047273.1; PID:g1518051
 C:Genetics:
 A:Gene: dfz2
 A:Cross-references: Flybase:FBgn0016797
 C:Superfamily: fruit fly frizzled protein
 C:Keywords: transmembrane protein

Query Match 48.9%; Score 43; DB 2; Length 694;
 Best Local Similarity 52.6%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHOESTXGRSGRSGXSG 19
 Db 182 SYTEAGSGSGSGSGSGSGS 200

RESULT 11
 H86209
 protein F22G5.10 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86209
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: H86209
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-843 <STO>
 A:Cross-references: GB:AE005172; NID:g8778569; PIDN:AAF79577.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F22G5.10
 A:Map position: 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.22396 Seconds

(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-9

Sequence: 1 SHQESTXGRSGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	88.6	416	1	P20930 homo sapien
2	52	59.1	506	1	P22420 human papil
3	44	50.0	774	1	WER2_MOUSE
4	43	48.9	694	1	FR22_DROME
5	43	48.9	1015	1	FRQ_CRESP
6	42	47.7	176	1	SSB_TREPA
7	42	47.7	308	1	GRP3_ARTSA
8	42	47.7	315	1	SOI2_YEAST
9	42	47.7	772	1	MR11_YEAST
10	42	47.7	2233	1	COAC_YEAST
11	41	46.6	825	1	SE5_RAT
12	40.5	46.0	379	1	ROA3_MOUSE
13	40	45.5	255	1	THO4_HUMAN
14	40	45.5	257	1	THO4_HUMAN
15	40	45.5	380	1	PEXD_PICPA
16	40	45.5	526	1	ZABR_YEAST
17	40	45.5	526	1	FUS_HUMAN
18	40	45.5	854	1	XR32_YEAST
19	39.5	44.9	135	1	RBPI_DROME
20	39.5	44.9	1895	1	WR19_ARATH
21	39.5	44.9	2594	1	7LES_DROVT
22	39	44.3	70	1	L2WU_ADE40
23	39	44.3	271	1	HEXP_LEIMA
24	39	44.3	419	1	TCF7_MOUSE
25	39	44.3	617	1	DBP1_YEAST
26	39	44.3	686	1	BDP1_YEAST
27	39	44.3	978	1	PEX6_RAT
28	39	44.3	980	1	PEX6_HUMAN
29	39	44.3	1703	1	SNF2_YEAST
30	38.5	43.8	697	1	AN3_XENLA
31	38	43.2	168	1	DH1_MAIZE
32	38	43.2	452	1	VE2_HPV17
33	38	43.2	483	1	VE2_HPV14

ALIGNMENTS

RESULT 1	ID	FILE_HUMAN	STANDARD	PRT	416 AA.
AC	P20930				
DT	01-FEB-1991	(Rel. 17, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Flilagrin precursor (Fragment).				
GN	FLG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Mammalia; Euteria; Primates; Carnivora; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
BN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89296901; PubMed=2740331;				
RA	McKinley-Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,				
RA	Canilizaro L., Croce C.M., Huebner K., Lessin S.R., Steiner P.M.,				
RT	"Characterization of a cDNA clone encoding human flaggrin and				
RT	localization of the gene to chromosome region 1q21."				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).				
RN	[2]				
RP	CITRULLINATION.				
RX	MEDLINE=96374388; PubMed=8780679;				
RA	Senshu T., Kan S., Ogawa H., Manabe M., Asaga H.,				
RT	"Preferential delamination of Keratin K1 and filaggrin during the				
RT	terminal differentiation of human epidermis."				
RL	Biochem. Biophys. Res. Commun. 225:712-719(1996).				
CC	-I- FUNCTION: Aggregates Keratin intermediate filaments and promotes				
CC	disulfide-bond formation among the intermediate filaments during				
CC	terminal differentiation of mammalian epidermis.				
CC	-I- PTM: Filaggrin is initially synthesized as a large, insoluble,				
CC	highly phosphorylated precursor containing many tandem copies of				
CC	324 AA, which are not separated by "large linker". The precursor				
CC	is deposited as keratohyalin granules. During terminal				
CC	differentiation it is dephosphorylated and proteolytically				
CC	cleaved.				
CC	-I- PTM: Undergoes delamination of some arginine residues				
CC	(citrullination).				
CC	-I- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL; M24355; AA52454.1; -				
DR	PIR; A32947; A32947.				
DR	Genew; HGNC:3748; FLG.				
DR	MTM; I35940; -				
DR	GO; GO:0005882; C:intermediate filament; NAS.				
DR	GO; GO:0005198; F:structural molecule activity; NAS.				
DR	GO; GO:0007275; P:development; NAS.				
DR	InterPro; IPR003303; Filaggrin.				

```

DR PFam; PF03516; Filaggrin; 2.
DR PRINTS; PR00487; FILAGRIN.
KW Phosphorylation; Citrullination; Developmental protein.
FT NON TER
SQ SEQUENCE 416 AA; 44105 MW; DEBA3218BA043F32 CRC64;

Query Match 88.6%; Score 76; DB 1; Length 416;
Best Local Similarity 84.2%; Pred. No. 1.4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGXSGS 19
7 SHQSTXGRSGRSGRSGS 25

RESULT 2
VE2_HP47 STANDARD; PRT; 506 AA.
AC P22420;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Regulatory protein E2.
GN E2.
OS Human papillomavirus type 47.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10594;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90281611; PubMed=2162112;
RA Kiyono T., Adachi A., Ishidashi M.;
RT "Genome organization and taxonomic position of human papillomavirus
RT type 47 inferred from its DNA sequence.";
RL Virology 177:401-405(1990).
CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACCCNNNNNGT-3'), PRESENT
CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2RE'S POSITION
CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
CC REPLICATION.
CC -1- SUBUNIT: Binds DNA as a dimer.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32305; AAA46979.1; -
DR FIR; D35324; W2ML47.
DR HSSP; P03122; 2BOP.
DR InterPro; IPR000427; E2_C.
DR InterPro; IPR001866; E2_N.
DR InterPro; IPR009021; Viral_DNA_bd.
DR PFam; PF00511; E2_C; 1.
DR PFam; PF00508; E2_N; 1.
DR ProDom; PD000672; E2_C; 1.
DR ProDom; PD000678; E2_N; 1.
KW Early protein; Transcription regulation; Activator; DNA-binding;
KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BB725065 CRC64;

Query Match 59.1%; Score 52; DB 1; Length 506;
Best Local Similarity 52.6%; Pred. No. 0.33;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGXSGS 19

```

```

DB 342 SREGNTRGRGRGRGRAGS 360

RESULT 3
MRK2_MOUSE STANDARD; PRT; 774 AA.
ID MRK2_MOUSE
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
GN MRK2 OR EMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Embryo;
RX MEDLINE=93364122; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19.";
RL Mamm. Genome 4:401-403(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC MARK subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70764; CA50040.1; -
DR FIR; 148609; 148609.
DR HSSP; Q63450; 1A06.
DR MGD; MGI:99638; Mark2.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR PFam; PF02149; KAI; 1.
DR PFam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SMO0220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 53 304
FT NP_BIND 59 67
FT BINDING 82 82
FT ACT_SITE 175 175
FT ACT_SITE BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BR8DBF44343A CRC64;

Query Match 50.0%; Score 44; DB 1; Length 774;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRSGRSGRSGXSGS 19
614 SPFGHGGRGRAGS 628

RESULT 4
FRZ2_DROME

```


FT CONFLICT 417 417 S -> T (IN REF. 1).
 SQ SEQUENCE 694 AA; 75451 MW; 6C510F13CBAFB96 CRC64;
 Query Match 48.9%; Score 43; DB 1; Length 694;
 Best Local Similarity 52.6%; Pred. No. 14;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 SHQSTXGRSRGRSGSGS 19
 DB 182 SYTAGSGGGSGSGSGS 200

RESULT 5

FRQ_CRESP STANDARD; PRT; 1015 AA.

AC 000586;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Frequency clock protein.

GN PRO.
 OS Creopus spinulosus (Chromocrea spinulosa).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocremycetidae; Hypocreales; Hypocreaceae; Hypocrea.

OX NCBI_TaxID=110619;
 RN [1]
 RC SEQUENCE FROM N.A.

RA STRAIN=PGSC 4896;
 RX MEDLINE=97186515; Pubmed=9037100;

RT Lewis M.T., Morgan L.W., Feldman J.F.;
 "Analysis of frequency (frq) clock gene homologs: evidence for a
 helix-turn-helix transcription factor";

RT Mol. Gen. Genet. 253:401-414(1997).

CC -1- FUNCTION: Circadian clock component involved in the generation of
 biological rhythms, in particular in rhythm stability, period
 length, and temperature compensation. Behaves as a negative

CC element in circadian transcriptional loop (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FRQ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC DR EMBL: U25850; AAA68072.1; -.

CC PIR: T42013; T42013.

CC KM Biological rhythms; Transcription regulation; Nuclear protein.

FT DOMAIN 240 245 POLY-SER.

FT DOMAIN 356 368 POLY-GLN.

FT DOMAIN 443 451 POLY-PRO.

FT DOMAIN 584 588 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 892 913 ASP/GLU-RICH (ACIDIC).
 FT SEQUENCE 1015 AA; 110972 MW; EA49E732ED7414B1 CRC64;

Qy Query Match 48.9%; Score 43; DB 1; Length 1015;
 Best Local Similarity 42.1%; Pred. No. 21;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SHQSTXGRSRGRSGSGS 19

DB 583 NHRQKXGHRGSDSGSSGN 601

RESULT 6
 SSB TREPA STANDARD; PRT; 176 AA.

AC 083101;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

RT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Single-strand binding protein (SSB) (Helix-destabilizing protein).
 GN SSB OR TP0062.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=160;
 RN [1]
 RC SEQUENCE FROM N.A.

RA STRAIN=Nichols;
 RX MEDLINE=98332770; Pubmed=965876;

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 Sodergren E., Hardham U.M., McLeod M.P., Salzberg S., Peterson J.,
 Khairak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 Venter J.C.;

RA "Complete genome sequence of Treponema pallidum, the syphilis
 RT spirochete";

RL Science 281:375-388(1998).

CC -1- FUNCTION: This protein is essential for replication of the
 CC chromosome. It is also involved in DNA recombination and repair
 CC (By similarity).

CC -1- SIMILARITY: Contains 1 SSB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC DR EMBL: AE001191; AAC65057.1; -.

CC PIR: F71370; F71370.

CC DR HSSP: P02339; IIEYG.

CC DR TIGR: TP0062; -.

CC DR InterPro: IPR006994; Nucleic acid DB.

CC DR InterPro: IPR000424; SSB_protein.

CC DR Pfam: PF00435; SSB; 1.

CC DR TIGRFAMs: TIGR00621; ssb; 1.

CC DR PROSITE: PS50935; SSB; 1.

CC KM DNA-binding; DNA repair; DNA replication; Complete proteome.

FT DOMAIN 4 106 SSB.

FT SEQUENCE 176 AA; 19043 MW; 47C49D2F7927D2C CRC64;

Qy Query Match 47.7%; Score 42; DB 1; Length 176;
 Best Local Similarity 62.5%; Pred. No. 4.6;
 Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 ESTYGRSRGRSGSGS 19

DB 128 ESTAGRVRGTSDDRS 143

RESULT 7
 GRP3 ARTSA STANDARD; PRT; 308 AA.

AC P13330;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glycine-rich protein GRP33.

OS Artemia salina (Brine shrimp).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.

OX NCBI_TaxID=85549;

RN [1]
 RC SEQUENCE FROM N.A.

RA MEDLINE=88007550; Pubmed=2443491;

RA Cruz-Alvarez M., Pellicer A.;

RT "Cloning of a full-length complementary DNA for an Artemia salina
 RT glycine-rich protein. Structural relationship with RNA binding
 RT proteins";

RL J. Biol. Chem. 262:13377-13380(1987).
 CC -1- PTM: THE ARGININES IN THE GLY-RICH DOMAIN MIGHT BE METHYLATED.
 CC -1- SIMILARITY: TO HD40, THE MAJOR PROTEIN COMPONENT OF ARTEMIA
 CC HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN PARTICLES, AND
 CC STRUCTURALLY TO OTHER NUCLEAR RNA BINDING PROTEINS.
 CC -1- SIMILARITY: Contains 1 KH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03453; AAC93400.1; -.
 DR PIR: A29379; A29379.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR SMART: SM00322; KH_1.
 DR PROSITE: PS50084; KH_type_1; 1.
 DR Nuclear protein; Ribonucleoprotein; Methylation; RNA-binding.
 FT DOMAIN 83 118 KH.
 FT 186 308 GLY-RICH.
 SQ SEQUENCE 308 AA; 33034 MW; E534CE552BB3BC2 CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 308;
 Best Local Similarity 72.7%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 GSRGRSGXSG 18
 DB 211 GGRGRGSGSG 221
 RESULT 8
 SOL2_YEAST STANDARD; PRT; 315 AA.
 ID SOL2_YEAST
 AC F37262;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable 6-phosphogluconolactonase 2 (EC 3.1.1.31) (PGL2).
 GN SOL2 OR YCXX13W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RX MEDLINE=96363906; PubMed=8725220;
 RA Shen W.C., Stanford D.R., Hopper A.K.;
 RT "Lossip, involved in yeast pre-tRNA splicing, positively regulates
 RT members of the SOL gene family.";
 RL Genetics 143:699-712(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ballista J.F.G., Franco L., Hoenicka J., Jimenez A., Remacha M.,
 RA Sanz E.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION, AND SIMILARITY.
 RX MEDLINE=94147996; PubMed=8313894;
 RA Koonin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RL EMBO J. 13:493-503(1994).
 CC -1- FUNCTION: Hydrolysis of 6-phosphogluconolactone to 6-
 CC phosphogluconate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
 CC phospho-D-gluconate.
 CC -1- PATHWAY: Pentose phosphate pathway; second step.
 CC -1- SIMILARITY: Belongs to the glucosamine/galactosamine-6-phosphate
 CC isomerase family. 6-phosphogluconolactonase subfamily.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U46559; AAB49322.1; -.
 DR EMBL: X59720; CAA42272.1; -.
 DR PIR: S53589; S53589.
 DR Germonline: 139004; -.
 DR SGD: S0000718; SOL2.
 DR GO: GO:0008033; P:RNA processing; IGI.
 DR InterPro: IPR006148; Gluc_gal_isom.
 DR InterPro: IPR005900; Phosphogluconlac.
 DR Pfam: PF01182; Glucosamine Iso; 1.
 DR TIGRFAMs: TIGR01198; p91, 1.
 DR Hydrolyase.
 FT CONFLICT 171 171 A -> V (IN REF. 2).
 FT SEQUENCE 315 AA; 34501 MW; 7F0AD76574AB276 CRC64;
 Query Match 47.7%; Score 42; DB 1; Length 315;
 Best Local Similarity 50.0%; Pred. No. 8.7;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 4 ESTYGRSGRSGXSGS 19
 DB 60 KSTASAEKSGSGSGS 75
 RESULT 9
 NR1_CAEEL STANDARD; PRT; 772 AA.
 ID NR1_CAEEL
 AC Q23255;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Double-strand break repair protein mre-11.
 GN MRE-11 OR ZC302.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RX MEDLINE=21135651; PubMed=11238374;
 RA Chin G.M., Villeneuve A.M.;
 RT "C. elegans mre-11 is required for meiotic recombination and DNA
 RT repair but is dispensable for the meiotic G(2) DNA damage
 RT checkpoint.";
 RL Genes Dev. 15:522-534(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kelly P.F.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR).
 CC Possesses single-strand endonuclease activity and double-strand-
 CC specific 3',5' exonuclease activity. Also involved in meiotic DSB
 CC processing.
 CC -1- FUNCTION: Required for meiotic crossing over and chiasma
 CC formation. Pachytene morphology and homolog pairing are normal.
 CC Vital in long term for maintenance of reproductive capacity of
 CC subsequent generations.
 CC -1- CORFACTOR: Manganese (By similarity).
 CC -1- SUBUNIT: Forms a complex with rad-50 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the NR1/RAD32 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z73978; CAA96292.1; -
 DR PIR; T27512; T27512.
 DR GenBank; 208612; -
 DR WormPep; ZC302.1; CE06573.
 DR InterPro; IPR003701; DNA_repair.
 DR InterPro; IPR004843; M-peptidase.
 DR InterPro; IPR007281; Mre11_DNA_bind.
 DR Pfam; PF00149; Metallophos; 1.
 DR Pfam; PF04152; Mre11_DNA_bind; 1.
 DR TIGRFAMs; TIGR00583; mre11; 1.
 DR DNA_repair; Hydrolase; Nuclease; Endonuclease; Exonuclease;
 KM Nuclear protein; Manganese; Melosis.
 FT MDTAGEN 139
 FT E-SK: IN MRE-11-ME41; DEFECTIVE IN
 FT MEIOTIC CHROMOSOME DEGRADATION.
 SQ SEQUENCE 772 AA; 86813 MW; 744A0754C260AC4B CRC64;

Query Match 47.7%; Score 42; DB 1; Length 772;
 Best Local Similarity 47.4%; Pred. No. 23;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 SHGRTXGRSGRSGXSGS 19
 Db 697 SKQPTTRGRGRGARGAGAS 715

RESULT 10
 COAC_YEAST STANDARD; PRT; 2233 AA.
 AC 000955;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [Includes: Biotin
 DE carboxylase (EC 6.3.4.14)].
 GN FAS3 OR ACC1 OR YNR016C OR N3175.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 2015-2022.
 RX MEDLINE=92262474; PubMed=1350093;
 RA Al-Feel W., Chirala S.S., Wakil S.J.;
 RT "Cloning of the yeast FAS3 gene and primary structure of yeast
 RT acetyl-CoA carboxylase";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4534-4538(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Pohl T.M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBD databases.
 CC -1- FUNCTION: This protein carries three functions: biotin carboxyl
 CC carrier protein, biotin carboxylase, and carboxyltransferase.
 CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate
 CC + malonyl-CoA.
 CC -1- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
 CC = ADP + phosphate + carboxybiotin-carrier protein.
 CC -1- COFACTOR: Biotin.
 CC -1- ENZYME REGULATION: By phosphorylation.
 CC -1- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
 CC step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: PARTIAL TO CARBAMOYL PHOSPHATE SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M92156; AAA20073.1; -
 DR EMBL; Z71631; CAA96294.1; -
 DR PIR; S63347; S63347.
 DR HSSP; P24182; IENC.
 DR GenBank; 143361; -
 DR SGD; S0005299; ACC1.
 DR GO; GO:0005789; Cytoplasmic reticulum membrane; IDA.
 DR GO; GO:0003989; F:acetyl-CoA carboxylase activity; IMP.
 DR GO; GO:0004075; F:biotin carboxylase activity; IMP.
 DR InterPro; IPR001882; Biotin_BS.
 DR InterPro; IPR005482; Biotin_carb_C.
 DR InterPro; IPR000089; Biotin_lipoYL.
 DR InterPro; IPR000022; Carboxyl_trans.
 DR InterPro; IPR005479; Case_L_D2.
 DR InterPro; IPR005481; Case_L_N.
 DR Pfam; PF02785; Biotin_carb_C; 1.
 DR Pfam; PF00364; Biotin_lipoYL; 1.
 DR Pfam; PF01039; Carboxyl_trans; 1.
 DR Pfam; PF00289; CPase_L_D2; 1.
 DR Pfam; PF02786; CPase_L_D2; 1.
 DR PROSITE; PS00188; BIOTIN; 1.
 DR PROSITE; PS00866; CPASE_1; 1.
 DR PROSITE; PS00867; CPASE_2; 1.
 KM Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme;
 KW ATP-binding; Phosphorylation
 FT NP_BIND 256 261
 FT ACT_SITE 383 383
 FT BINDING 735 735
 FT DOMAIN 1865 1894
 FT CONFLICT 1523 1523
 FT CONFLICT 1755 1755
 FT CONFLICT 1761 1766
 SQ SEQUENCE 2233 AA; 250351 MW; 0A335AAD931P8308 CRC64;

Query Match 47.7%; Score 42; DB 1; Length 2233;
 Best Local Similarity 50.0%; Pred. No. 72;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HOESTXGRSGRSGXSGS 19
 Db 1200 HQSSNGPAPDRSGSAS 1217

RESULT 11
 S55_RAT STANDARD; PRT; 825 AA.
 AC Q6303;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE S55 antigen.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistar; TISSUE=Brain;
 RX MEDLINE=96015159; PubMed=8537300;
 RA Suzuki E., Kojima N., Yoshimura K., Uyemura K., Obata K., Akagawa K.;
 RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
 RT protein S55 in the nervous system";
 RL J. Biochem. 118:122-128(1995).
 CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Expressed in neurons.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D37934; BAA07153.1; -;
 DR PIR; JC4163; JC4163.
 KW DNA-binding; Nuclear protein; Antigen.
 SQ SEQUENCE 825 AA; 86831 MW; AF667FE2FD55BDF CRC64;

Query Match 46.5%; Score 41; DB 1; Length 825;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GRSRGRSGSGS 19
 Db 589 GRGRGRGRSGS 600

RESULT 12
 ID ROA3_MOUSE STANDARD; PRT; 379 AA.
 AC Q8BG05; Q8BHF8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Heterogeneous nuclear ribonucleoprotein A3 (hnRNP A3).
 GN HNRPA3 OR HNRNP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Bayarsaahan D.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=FVB/N; TISSUE=Eye;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish P.,
 RA Diatchenko L., Mansura K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebai T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shewchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 RL Proc Natl Acad Sci U S A. 99:16899-16903(2002).
 CC -1- SUBCELLULAR LOCATION: Nuclear; component of ribonucleosomes (By
 CC similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8BG05-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8BG05-2; Sequence=VSP_007350;
 CC -1- SIMILARITY: BELONGS TO THE A/B GROUP OF HNRNP, WHICH ARE BASIC AND
 CC GYR-RICH PROTEINS.

CC -1- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF463524; AAN7692.1; -;
 DR EMBL; BC023828; AAH23828.1; -;
 DR EMBL; BC023908; AAH23908.1; -;
 DR EMBL; BC038364; AAH38364.1; -;
 DR InterPro: IPR005054; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 2.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS50102; RRM; 2.
 DR PROSITE; PS00030; RRM_RNP_1; 2.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
 KW Alternative splicing.
 FT DOMAIN 35 118 RNA-BINDING (RRM) 1.
 FT DOMAIN 126 205 RNA-BINDING (RRM) 2.
 FT DOMAIN 211 379 GYR-RICH
 FT VARSPLIC 1 23 MEKPPGRPOPDSGRRRRRGE -> M (in isoform
 FT 2).
 SQ SEQUENCE 379 AA; 39652 MW; D83C400A2B096E9B CRC64;

Query Match 46.0%; Score 40.5; DB 1; Length 379;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 SHQE-STXGRSGRSGSGS 19
 Db 203 SKQEWQSGSGRGRSGSGN 222

RESULT 13
 ID THO4_MOUSE STANDARD; PRT; 255 AA.
 AC Q08583; Q8CBM4; Q9JW7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE THO complex subunit 4 (Tho4) (RNA and export factor binding protein 1)
 DE (REF1-1) (Ally of AMU-1 and LEF-1) (Ally/REF).
 GN THO4 OR REF1P OR REF1 OR ALY.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR LOCATION, AND
 RP TISSUE SPECIFICITY.
 RC TISSUE=Thymocytes;
 RX MEDLINE=972030366; PubMed=9119228;
 RA Brum L., Munerlyn A., Grosschedl R.,
 RT "Ally, a context-dependent coactivator of LEF-1 and AMU-1, is required
 RT for TCRAlph enhancer function."
 RL Genes Dev. 11:640-653(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, AND INTERACTION WITH NXF1.
 RC TISSUE=Embryo;
 RX MEDLINE=20246526; PubMed=10786854;
 RA Stutz F., Bach A., Doerks T., Braun I.C., Seraphin B., Wilm M.,
 RA Bork P., Izaurralde E.,
 RT "REF, an evolutionarily conserved family of hnRNP-like proteins,
 RT interacts with TAP/Mex57p and participates in mRNA nuclear export."
 RL RNA 6:638-650(2000).
 RN [3]
 RP SEQUENCE OF 190-255 FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Urinary bladder;

RA MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikida I., Otsu N., Saito R., Suzuki H., Yamahara I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gotohori T.,
 RA Bardelli R., Hill D.P., Bult C., Hume D.A., Quakenbush J.,
 RA Schiml L.M., Karapin A., Matsuda H., Batatov S., Betsel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazier K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Katsuya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numa K., Okito T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid U., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Sempke C.A., Secou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verrardo R., Wagner L., Wahlestedt C., Wang Y., Waterman J., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Zimin P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shitaki T., Waki K., Kawai U., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RT Nature 420:563-573 (2002).
 [4]
 RA INTERACTION WITH THE EXON JUNCTION COMPLEX.
 RA MEDLINE=22088306; PubMed=12093754;
 RA Lejeune F., Ishigaki Y., Li X., Maquat L.E.;
 RT "The exon junction complex is detected on CBP80-bound but not
 RT eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling.";
 RT EMBO J. 21:3536-3545 (2002).
 [5]
 RA STRUCTURE BY NMR OF 105-182.
 RA MEDLINE=22694206; PubMed=12803490;
 RA Perez-Alvarado G.C., Martinez-Yamout M., Allen M.M., Groschedl R.,
 RA Dyson H.U., Wright P.E.;
 RT "Structure of the nuclear factor ALY: insights into
 RT post-transcriptional regulatory and mRNA nuclear export processes.";
 RT Biochemistry 42:7348-7357 (2003).
 -1- FUNCTION: Acts as a chaperone and promotes the dimerization of
 CC transcription factors containing basic leucine zipper (bZIP)
 CC domains and thereby promotes transcriptional activation. Plays a
 CC role in mRNA processing and export. May function as a scaffold that
 CC mediates interactions between proteins and/or RNA. Integral part
 CC of the THO/TEX complex that is recruited to transcribed genes and
 CC travels with the RNA polymerase during elongation. Is part of the
 CC exon junction complex that remains associated with spliced mRNA
 CC and plays an important role in mRNA export and nonsense-mediated
 CC RNA decay.
 -1- SUBUNIT: Homomultimer. Binds LEF1 and RUNX1. Is part of several
 CC complexes involved in mRNA processing and export. Part of the
 CC heteromultimeric THO/TEX complex containing THOC1, THOC2, THOC3,
 CC THOC4 and NXF1/VAPB6. Associates with the spliceosome (BY
 CC similarity). Binds NXF1, NXF2 and RBM8A directly and is part of
 CC the exon junction complex (EJC) containing NCBP1, NCBP2, NCBP3,
 CC RBM8A, SRRM1, NXF1, NXF2, UPF3B, UPF2, THOC4 and/or RBM8A.
 -1- SUBCELLULAR LOCATION: Nuclear. Travels to the cytoplasm as part of
 CC the exon junction complex (EJC) bound to mRNA.
 -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=RefSeq-1;
 CC IsoId=008583-1; Sequence=Displayed;
 CC Name=2; Synonyms=RefSeq-11;
 CC IsoId=008583-2; Sequence=VSP_008597;
 -1- TISSUE SPECIFICITY: Highly expressed in heart, brain, spleen,
 CC lung, liver, skeletal muscle, kidney and testis.
 -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (see <http://www.isb-stb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 DR EMBL; U89876; AAC53117.1; -;
 DR EMBL; AJ252140; CAB76383.1; -;
 DR EMBL; AK035721; BAC29168.1; -;
 DR PDB; 1N08; 12-AUG-03.
 DR TRANSFAC; T03213; -;
 DR MGD; MGI:134104; RefSeq.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NRG.
 KW Chapterone; Transport; mRNA transport; mRNA processing; mRNA splicing;
 KW Nuclear protein; Spliceosome; RNA-binding; Alternative splicing;
 KW 3D-structure.
 FT DOMAIN 21 229 GUY/ALA/ARG-RICH.
 FT DOMAIN 105 182 RNA-BINDING (RRM).
 FT VARSPIC 14 105 Missing (in isoform 2).
 FT /FTID=VSP_008597.
 FT SEQUENCE 255 AA; 26940 MW; F597235EBDD47C17 CEC64;
 Query Match 45.5%; Score 40; DB 1; Length 255;
 Best Local Similarity 63.6%; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 8 GRSRGRSGXSG 18
 DB 26 GRSRGRSGXSG 36
 RESULT 14
 ID THO4_HUMAN STANDARD; PRT; 257 AA.
 AC Q86V81; Q43672;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE THO complex subunit 4 (Tho4) (Ally of AML-1 and LRF-1)
 DE (Transcriptional coactivator Aly/REF) (bZIP enhancing factor BRF).
 GN THO4 OR Aly OR BRF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marnusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Wilting M., Madan A., Young A.C., Shvetchenko I., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT [2].
RP SEQUENCE OF 25-257 FROM N.A., AND DISEASE.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=99135498; PubMed=9952027;
RA Wichmann I., Garcia-Lozano J.R., Respaldaiza N.,
RA Gonzalez-Escribano M.F., Nunez-Roldan A.;
RT "Antibodies to transcriptional regulation proteins DEK and ALY in
RT a patient with systemic lupus erythematosus.";
RL Hum. Immunol. 60:57-62(1999).
RN [3].
RP SEQUENCE OF 127-111, 182-189, 239-245 AND 249-251, FUNCTION, AND
RP SUBCELLULAR LOCATION.
RX MEDLINE=99417961; PubMed=10488337;
RA Virdasius C.-M., Wagner S., Green M.R.;
RT "A human nuclear-localized chaperone that regulates dimerization, DNA
RT binding, and transcriptional activity of bZIP proteins.";
RL Mol. Cell 4:219-228(1999).
RN [4].
RP PARTIAL SEQUENCE, MASS SPECTROMETRY, AND INTERACTION WITH THE
RP SPLICEOSOME.
RX MEDLINE=98400255; PubMed=9731529;
RA Neubauer G., King A., Rappalber J., Calvio C., Watson M., Ajuh P.,
RA Sleeman J., Lamond A., Mann M.;
RT "Mass spectrometry and EST-database searching allows characterization
RT of the multi-protein spliceosome complex.";
RL Nat. Genet. 20:46-50(1998).
RN [5].
RP SEQUENCE OF 108-133, MASS SPECTROMETRY, AND INTERACTION WITH THE
RP SPLICEOSOME.
RX MEDLINE=22166132; PubMed=12176931;
RA Rappalber J., Ryder U., Lamond A.I., Mann M.;
RT "Large-scale proteomic analysis of the human spliceosome.";
RL Genome Res. 12:1231-1245(2002).
RN [6].
RP FUNCTION, AND INTERACTION WITH RBM8A; NXF1 AND THE EXON JUNCTION
RP COMPLEX.
RX MEDLINE=21564074; PubMed=11707413;
RA Katakoka N., Diem M.D., Kim V.N., Yong J., Dreyfuss G.;
RT "Mago, a human homolog of Drosophila mago nashi protein, is a
RT component of the splicing-dependent exon-exon junction complex.";
RL EMBO J. 20:6424-6433(2001).
RN [7].
RP FUNCTION, AND INTERACTION WITH NXF1 AND NXF1.
RX MEDLINE=21530486; PubMed=11675789;
RA Luo M.-L., Zhou Z., Magni K., Christoforides C., Rappalber J.,
RA Mann M., Reed R.;
RT "Pre-mRNA splicing and mRNA export linked by direct interactions
RT between UAP56 and Aly.";
RL Nature 413:644-647(2001).
RN [8].
RP FUNCTION, AND INTERACTION WITH ULS4.
RX MEDLINE=22326840; PubMed=12438613;
RA Chen I.-H.B., Schlabach K.S., Sandri-Goldin R.M.;
RT "ICP27 interacts with the RNA export factor Aly/REF to direct herpes
RT simplex virus type 1 intronless mRNAs to the TAP export pathway.";
RL J. Virol. 76:12877-12889(2002).
RN [9].
RP FUNCTION, AND INTERACTION WITH THE TREX COMPLEX.
RX MEDLINE=22010388; PubMed=11979277;
RA Straesser K., Masuda S., Mason P., Pfannstiel J., Opijzki M.,
RA Rodriguez-Navarro S., Rondon A.G., Aguilera A., Struhl K., Reed R.,
RA Hurt E.;
RT "TREX is a conserved complex coupling transcription with messenger
RT RNA export.";
RL Nature 417:304-308(2002).
RN [10].
RP FUNCTION: Acts as chaperone and promotes the dimerization of
RP transcription factors containing basic leucine zipper (bZIP)
RP domains and thereby promotes transcriptional activation. Plays a
RP role in mRNA processing and export. May function as scaffold that
RP mediates interactions between proteins and/or RNA. Integral part

CC of the THO/TREX complex that is recruited to transcribed genes and
CC travels with the RNA polymerase during elongation. Is part of the
CC exon junction complex that remains associated with spliced mRNA
CC and plays an important role in mRNA export and nonsense-mediated
CC RNA decay. Directs mRNA derived from Herpes simplex virus intron-
CC less genes to the NXF1-mediated export pathway.
CC
CC -1- SUBUNIT: Homooligomer. Binds REF1 and RUV1 (by similarity). Is
CC part of several complexes involved in mRNA processing and export.
CC Part of the heteromultimeric THO/TREX complex containing THOC1,
CC THOC2, THOC3, THOC4 and NXF1/UAP56. Associates with the
CC spliceosome. Binds NXF1 and RBM8A directly and is part of
CC the exon junction complex (EJC) containing NCBP1, NCBP2, NBP1,
CC RBM8A, SRRM1, NXF1, NXF1, UPF3B, UPF2 and THOC4 (by similarity).
CC Binds ULS4 from Herpes simplex type 1 (strain KOS).
CC
CC -1- SUBCELLULAR LOCATION: Nuclear. Travels to the cytoplasm as part of
CC the exon junction complex (EJC) bound to mRNA.
CC
CC -1- DISEASE: Antibodies against THOC4 are found in sera of patients
CC with systemic lupus erythematosus (SLE).
CC
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL, BC052302; AAH52302.1; -
CC EMBL, AF047002; AAD09608.1; -
CC Genew; HGNC:19071; THOC4.
CC MIM; 604171; -
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rtm; 1.
CC
CC SMART; SM00360; RNM; 1.
CC
CC PROSITE; PS50102; RNM_1; FALSE NEG.
CC
CC DR PROSITE; PS00030; RNM; 1.
CC
CC DR Antigen; Chaperone; Transport; mRNA transport; mRNA processing;
CC
CC KW mRNA splicing; Nuclear protein; Spliceosome; RNA-binding.
CC
CC FT DOMAIN 21 231
CC FT 106 183 RNM-BINDING (RNM).
CC FT 26 26 G->R (IN REF. 2).
CC FT 35 36 G->R (IN REF. 2).
CC FT 39 39 D->N (IN REF. 2).
CC FT 150 150 D->N (IN REF. 2).
CC FT 169 169 V->F (IN REF. 2).
CC
CC FT CONFLICT 169 169
CC
CC SQ SEQUENCE 257 AA; 26888 MW; E2B5021DA579919A CRC64;
CC
CC
CC Query Match 45.5%; Score 40; DB 1; Length 257;
CC Best Local Similarity 63.6%; Pred. No. 15;
CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC
CC QY 8 GSRGRSGXSG 18
CC Db 26 GRGRGRASQG 36
CC
CC
CC RESULT 15
CC PEXD_PICPA STANDARD; PRT; 380 AA.
CC AC 092266;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Peroxisomal membrane protein PEX13 (Peroxin-13).
CC GN PEX13.
CC OS Pichia pastoris (Yeast).
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC OC Saccharomycetales; Saccharomycetaceae; Pichia.
CC OX NCB:taxid=4922;
CC RN (1)
CC RP SEQUENCE FROM N.A., AND MUTAGENESIS.
CC MEDLINE=97011155; PubMed=8858165;
CC

```

RA Gould S.J., Kalish J.E., Morrell J.C., Bjorkman J., Urquhart A.J.,
RA Crane D.I.;
RT "Pex13p is an SH3 protein of the peroxisome membrane and a docking
RT factor for the predominantly cytoplasmic Pts1 receptor.";
RL J. Cell Biol. 135:85-95(1996).
CC -!- FUNCTION: Component of the peroxisomal translocation machinery
CC with PEX14 and PEX17. Interacts with the Pts1 receptor
CC (PAG10/PEX5). Involved in the import of Pts1 and Pts2 proteins.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, PEX13L (shown here) and PEX13S, are produced
CC by alternative initiation;
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U70067; AAB09087.1; -.
DR HSSP; Q64010; ICKA.
DR InterPro; IPR007223; Peroxin-13_N.
DR InterPro; IPR001452; SH3.
DR Pfam; PF04088; Peroxin-13_N; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Peroxisome; Transport; Protein transport; Transmembrane; SH3 domain;
KW Alternative initiation.
FT CHAIN 1 380
FT FT 33 380
FT FT 33 380
FT FT 33 380
FT INIT MET 33 33
FT DOMAIN 1 230
FT TRANSMEM 231 251
FT DOMAIN 252 380
FT DOMAIN 277 344
FT MUTAGEN 286 286
FT MUTAGEN 287 287
FT MUTAGEN 288 288
FT MUTAGEN 291 291
FT MUTAGEN 296 296
SQ SEQUENCE 380 AA; 40695 MW; EEBAC39F93BA832 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 380;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 SHOESTXGRSGRSGXSG 18
DB 315 SNOESTYWWKCRSRDXYG 332

```

Search completed: September 28, 2004, 06:05:52
 Job time : 8.22396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 40.1771 Seconds
(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-9

Perfect score: 88
Sequence: 1 SHQESTXGRSGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL 25:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_virussel:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	95.5	797	4	Q16824
2	80	90.9	990	4	Q15206
3	80	90.9	1218	4	Q05331
4	78	88.6	798	4	Q9H4U3
5	78	88.6	1084	4	Q01212
6	69	78.4	591	4	Q01720
7	64	72.7	687	4	Q9H4U2
8	63	71.6	322	4	Q03838
9	63	71.6	465	4	Q03838
10	49	55.7	820	12	Q69375
11	49	55.7	822	3	Q9P3T2
12	47	53.4	436	16	Q82C67
13	46	52.3	476	12	Q80890
14	46	52.3	810	10	Q94B50
15	46	52.3	1142	13	Q7SZV2
16	44	50.0	285	5	Q9W3V3

17	44	50.0	691	4	Q96RC0	Q96RC0 homo sapien
18	44	50.0	722	11	Q08679	Q08679 rattus norv
19	44	50.0	745	4	Q15524	Q15524 homo sapien
20	44	50.0	755	4	Q96HB3	Q96HB3 homo sapien
21	44	50.0	787	5	Q9GSD4	Q9GSD4 plasmodium
22	44	50.0	888	11	Q8BR95	Q8BR95 mus muscula
23	44	50.0	936	16	Q82H71	Q82H71 streptomyce
24	44	50.0	1284	5	Q9V9Y3	Q9V9Y3 caenorhabdi
25	44	50.0	1656	5	Q21948	Q21948 drosophila
26	44	50.0	1765	5	Q8SX63	Q8SX63 drosophila
27	44	50.0	2232	5	Q9VC96	Q9VC96 drosophila
28	43	48.9	306	5	Q45396	Q45396 caenorhabdi
29	43	48.9	611	10	Q9S767	Q9S767 oryza sativ
30	43	48.9	652	5	Q77168	Q77168 apis mellif
31	43	48.9	756	5	Q9U116	Q9U116 drosophila
32	43	48.9	802	10	Q8S2A7	Q8S2A7 oryza sativ
33	43	48.9	813	10	Q8VZ18	Q8VZ18 arabidopsis
34	43	48.9	843	10	Q9LNX5	Q9LNX5 arabidopsis
35	43	48.9	849	10	Q9C7M2	Q9C7M2 arabidopsis
36	43	48.9	849	10	Q949M6	Q949M6 arabidopsis
37	42.5	48.3	1080	12	Q91AM0	Q91AM0 beet chloro
38	42	47.7	188	10	Q8H8P6	Q8H8P6 oryza sativ
39	42	47.7	199	10	Q9FW52	Q9FW52 oryza sativ
40	42	47.7	240	16	Q87YL4	Q87YL4 pseudomonas
41	42	47.7	297	6	Q8S024	Q8S024 bos taurus
42	42	47.7	322	16	Q8G4E5	Q8G4E5 bifidobacte
43	42	47.7	336	10	Q943T8	Q943T8 oryza sativ
44	42	47.7	481	16	Q7UJP9	Q7UJP9 rhodospirell
45	42	47.7	507	2	Q9RMT5	Q9RMT5 streptomyce

ALIGNMENTS

Q16824	PRELIMINARY;	PRT;	797 AA.
AC Q16824;			
DT 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE Profilaggrin (Fragment).			
GN FLG.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=91064347; PubMed=2248957;			
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Seimert P.M.,			
RT "Organization, structure, and polymorphisms of the human profilaggrin			
RT gene [published erratum appears in Biochemistry 1991 Jun			
RT 11:30(23):5814]."			
RL Biochemistry 29:9432-9440(1990).			
DR EMBL; M60502; AA663248.1;			
DR GO; GO:0005198; F:structural molecule activity; IEA.			
DR InterPro; IPR003303; Filaggrin.			
DR Pfam; PF03516; Filaggrin; 4.			
DR PRINTS; PRO0487; FILAGRIN.			
FT NON TER			
SQ SEQUENCE 797 AA; 85176 MW; 6056184763BDA86B CRC64;			
QY	1	SHQESTXGRSGRSGXSGS 19	95.5%; Score 84; DB 4; Length 797;
DB	427	SHQESTXGRSGRSGRSGS 445	Best Local Similarity 89.5%; Pred. No. 1.3e-05;
			Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 2

```

O15206
ID Q15206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
PR Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene [published extratum appears in Biochemistry 1991 Jun
11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Placenta;
RX MEDLINE=91255199; PubMed=2043621;
RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
gene.";
RL Biochemistry 30:5814-5814(1991).
DR EMBL; M60494; AA63244.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
FT NON TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 90.9%; Score 80; DB 4; Length 990;
Best Local Similarity 84.2%; Pred. No. 7.8e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
DB 551 SHOESTGRSGRSGSGS 569

RESULT 3
Q05331 PRELIMINARY; PRT; 1218 AA.
AC Q05331;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
PR FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.-Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-625(1993).
DE - FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS
CC - PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC CC -1- PROTEOLYTICALLY CLEAVED.
CC CC POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC CC REPEATS.
DR EMBL; M96943; AA36487.1; -
DR PIR; A48118; A48118.
DR HSP; P02593; 1CDM.
DR GO; GO:0005856; C:cytoskeleton; NAS.
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR GO; GO:0030154; P:cell differentiation; NAS.
DR GO; GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro; IPR001751; CAPB S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF03516; Filaggrin; 6.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CAPB; 1.
KM Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KM Polymorphism.
FT CA BIND 19 32 SITE I (BY SIMILARITY).
FT CA BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 1218 1218
SQ SEQUENCE 1218 AA; 133604 MW; EC195AD5285B19C2 CRC64;

Query Match 90.9%; Score 80; DB 4; Length 1218;
Best Local Similarity 84.2%; Pred. No. 9.8e-05;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
DB 773 SHOESTGRSGRSGSGS 791

RESULT 4
Q094U3 PRELIMINARY; PRT; 798 AA.
AC Q094U3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DJ14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356504; CAC13171.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 4.
DR PRINTS; PR00487; Filaggrin.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 88.6%; Score 78; DB 4; Length 798;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
DB 428 SHOESTGRSGRSGSGS 446

RESULT 5
Q01212 PRELIMINARY; PRT; 1084 AA.

```

```

AC Q01212; Q03840;
DR 01-NOV-1996 (TREMBLrel. 01, Created)
DR 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DR 01-UN-2003 (TREMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published structure appears in Biochemistry 1991 Jun
RT 11;30(23):5814].".
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60503; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NAS.
DR GO; GO:0005198; F:structural molecule activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin.
DR PRINTS; PR00487; Filaggrin.
DR NON TER
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 88.6%; Score 78; DB 4; Length 1084;
Best Local Similarity 84.2%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

CY 1 SHOE5TXGRSGRSGSGS 19
DB 64 SHOE5TXGRSGRSGSGS 82
|||||
Q01720 PRELIMINARY; PRT; 591 AA.
ID Q01720;
AC Q01720;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=93054736; PubMed=1429717;
RA Presland R.B., Haydock P.V., Fleckman P., Ntunusuksi W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus.".
RL J. Biol. Chem. 267:23772-23781(1992).
DR J. Biol. Chem. 267:23772-23781(1992).
CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
CC DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PFM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL; L01089; AAA60177.1; -.
DR EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467 POTENTIAL.
FT PROPEP 468 474 FILAGGRIN.
FT CHAIN 475 >591 POTENTIAL.
FT CA_BIND 19 32 FILAGGRIN.
FT CA_BIND 62 73 SITE I (BY SIMILARITY).
FT NON TER 591 591 SITE II (BY SIMILARITY).
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 591;
Best Local Similarity 78.9%; Pred. No. 0.0032;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 1 SHOE5TXGRSGRSGSGS 19
DB 449 SHOE5TXGRSGRSGSGS 467
|||||
Q094U2 PRELIMINARY; PRT; 687 AA.
ID Q094U2;
AC Q094U2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE DJ14N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laid G.;
RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; eHand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 78.4%; Score 69; DB 4; Length 687;
Best Local Similarity 78.9%; Pred. No. 0.0037;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 SHOESTXGRSGRSGSGS 19
 DB 449 SHOESTXGRSGRSGSGS 467

RESULT 8

ID 003838 PRELIMINARY; PRT; 465 AA.

AC 003838;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE FILAGGRIN (PROFILAGGRIN) (Fragment).
 GN FLG.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OK NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;
 RX MEDLINE=91064347; PubMed=2248957;
 RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene."

RL Biochemistry 29:9432-9440(1990).

RN [2]

REVISIONS.
 RX MEDLINE=91255199; PubMed=2043621;

RA Gan S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;

RT "Organization, structure, and polymorphisms of the human profilaggrin gene."

RL Biochemistry 30:5814-5814(1991).

CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.

CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN REPEATS.

CC -1- MISCELLANEOUS: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 324 AA. THE PRECURSOR IS DEPOSED AS KERATOHYALIN GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

DR EMBL; M60499; AAA63246.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 3.

DR PRINTS; PRO0487; FILAGGRIN.

FT NON_TER 1 1
 FT 465 465

SQ SEQUENCE 465 AA; 50280 MW; C883744C5B134097 CRC64;

QY 1 SHOESTXGRSGRSGSGS 19
 DB 227 SHOESTXGRSGRSGSGS 245

Query Match 72.7%; Score 64; DB 4; Length 465;
 Best Local Similarity 73.7%; Pred. No. 0.017;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 9

ID 075370 PRELIMINARY; PRT; 322 AA.

AC 075370;
 DT 01-NOV-1998 (TRENBLREL. 08, Created)

DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Epidermal filaggrin (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=99101527; PubMed=9886436;
 RA Girbal-Neuhauser E., Durieux J.J., Arnaud M., Dalbon P., Sebbaq M., Vincent C., Simon M., Sensu T., Masson-Bessiere C., RA Jolivet-Reyraud C., Jolivet M., Sere G.;

RT "The epitopes targeted by the rheumatoid arthritis-associated antifilaggrin autoantibodies are posttranslationally generated on various sites of (pro)filaggrin by deamination of arginine residues."

RL J. Immunol. 162:585-594(1999).

DR EMBL; AF043380; AAC23559.1; -;
 DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.

DR PRINTS; PRO0487; FILAGGRIN.

FT NON_TER 1 1
 FT 322 322

SQ SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FP9E0 CRC64;

QY 1 SHOESTXGRSGRSGSGS 18
 DB 305 SHOESTXGRSGRSGSGS 322

Query Match 71.6%; Score 63; DB 4; Length 322;
 Best Local Similarity 72.2%; Pred. No. 0.017;
 Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 10

ID 069375 PRELIMINARY; PRT; 820 AA.

AC 069375;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Putative U125 protein.

OS Mouse cytomegalovirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.

OK NCBI_TaxID=10366;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=K181;

RA MEDLINE=94233727; PubMed=7513920;
 RA Dallas P.B., Lyons P.A., Hudson J.B., Scalzo A.A., Shellam G.R.;

RT "Identification and characterization of a murine cytomegalovirus gene with homology to the U125 open reading frame of human cytomegalovirus."

RT Virology 200:643-650(1994).

DR EMBL; U02500; AAA19449.1; -;
 DR InterPro; IPR006731; Herpes_pp85.

DR Pfam; PF04637; Herpes_pp85; 1.

SQ SEQUENCE 820 AA; 90346 MW; 53638A232334F79C CRC64;

RESULT 11

ID 099312 PRELIMINARY; PRT; 822 AA.

AC 099312;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.

GN B12P1.10.

```

OS Neurospora crassa.
CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
CX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Miyakura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390091; CAB98213.1; -
DR PIR; T51049; T51049.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR005054; RNA rec mot.
DR InterPro; IPR005120; Smg-4_UPF3.
DR Pfam; PF00076; itm; 1.
DR Pfam; PF03467; Smg4_UPF3; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 55.7%; Score 49; DB 3; Length 822;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0.

OY 3 QESTXGRSGRSGXSGS 19
DB 414 RESASGRTRGRGRGCT 430

RESULT 12
O82C67 PRELIMINARY; PRT; 436 AA.
ID AC Q82C67;
AC Q82C67;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DS Putative integral membrane protein.
GN SAV5487.
OS Streptomyces avermitilis.
CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
CC Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948;
RX Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.,
RA "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005043; BAC73199.1; -
DR InterPro; IPR002194; Chaparomlin_TCP-1.
DR PROSITE; PS00995; TCP1_3; 1.
KM Complete proteome.
SQ SEQUENCE 436 AA; 45359 MW; 35BD46EB0C7282D0 CRC64;

```

[illegible]

DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_centra.
 DR InterPro; IPR003960; AAA_sub.
 DR InterPro; IPR005936; Peptidase_F1sh.
 DR InterPro; IPR00642; Peptidase_M1.
 DR Pfam; PF00004; AAA; 1.
 DR Pfam; PF01434; Peptidase_M41; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR01241; F1sh_fam; 1.
 DR PROSITE; PS00674; AAA; 1.
 KW ATP-binding; Metalloprotease; Protease.
 SQ SEQUENCE 810 AA; 89317 MW; 22A0EC7ADEB6C18 CRC64;

Query Match 52.3%; Score 46; DB 10; Length 810;
 Best Local Similarity 75.0%; Pred. No. 33;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 8 GRSRGRSGSGS 19
 Db 429 GRKRGCGGSGS 440

RESULT 15
 Q7SZV2 PRELIMINARY; PRT; 1142 AA.
 AC Q7SZV2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Period 1.
 DE PERL.
 GN PERL.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N. A.
 RA Constance C.M., Sulton B., Taylor M., Green C.B.;
 RT "Xenopus laevis period 1.", EMBL/GenBank/DBJ databases.
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY340103; AAQ18143.1;
 SQ SEQUENCE 1142 AA; 123519 MW; 158ECB698DEB3E0B6 CRC64;

Query Match 52.3%; Score 46; DB 13; Length 1142;
 Best Local Similarity 58.8%; Pred. No. 47;
 Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Oy 1 SHQSTYGRSRGRSGS 17
 Db 805 SKGRSNGRRRRGKSGS 821

Search completed: September 28, 2004, 06:12:48
 Job time : 41.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 54.1302 Seconds
(without alignments)
99.176 Million cell updates/sec

Title: US-09-308-150-9
Perfect score: 88
Sequence: 1 SHQESTXCRGRGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.290a04:*
1: geneseqp19808:*
2: geneseqp19905:*
3: geneseqp20008:*
4: geneseqp20015:*
5: geneseqp20025:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	96.6	19	2	AAW61515 Peptide c
2	84	95.5	19	2	AAW61517 Peptide c
3	84	95.5	19	2	AAW61509 Peptide c
4	84	95.5	19	2	AAW61513 Peptide c
5	84	95.5	19	2	AAW61505 Peptide c
6	84	95.5	19	2	AAW61514 Peptide c
7	84	95.5	19	2	AAW61516 Peptide c
8	80	90.9	14	5	ABR7605 Novel hum
9	78	88.6	19	2	AAW61506 Peptide c
10	78	88.6	19	2	AAW61508 Peptide c
11	78	88.6	19	2	AAW61507 Peptide c
12	78	88.6	19	2	AAW61512 Peptide c
13	78	88.6	19	2	AAW61511 Peptide c
14	78	88.6	19	2	AAW61510 Peptide c
15	75	85.2	21	2	AAW61520 Peptide c
16	75	85.2	22	4	AAW61523 Peptide c
17	58	65.9	330	2	AAW61525 Human fil
18	58	65.9	330	2	AAW61526 Human fil
19	58	65.9	330	2	AAW61527 Human fil
20	54	61.4	330	2	AAW61529 Human fil
21	51	58.0	641	4	ABG19110 Novel hum
22	50	56.8	477	6	ABO07142 Novel hum
23	45	51.1	1711	4	AAW79819 Human pro
24	45	51.1	1951	4	AAW78835 Human pro
25	44	50.0	285	4	ABW69359 Drosophila

26	44	50.0	532	4	ABW67173 Human PAR
27	44	50.0	691	5	ABW69050 Human PAR
28	44	50.0	691	6	ABW69050 Human PAR
29	44	50.0	722	6	ABW69051 Human PAR
30	44	50.0	724	5	ABW69051 Human PAR
31	44	50.0	724	6	ABW69052 Human PAR
32	44	50.0	724	6	ABW69053 Human PAR
33	44	50.0	731	7	ABW69074 Human KPP
34	44	50.0	745	6	ABW69350 Human mic
35	44	50.0	745	6	ABW69370 Human mic
36	44	50.0	745	7	ABW69376 Human ser
37	44	50.0	768	4	AAW78349 Human pro
38	44	50.0	774	6	ABW69369 Serine/th
39	44	50.0	777	6	ABW69368 Human ser
40	44	50.0	780	6	ABW69367 Human ser
41	44	50.0	804	6	ABW69333 Human MDP
42	44	50.0	804	4	AAW79333 Human pro
43	44	50.0	1129	4	ABW68274 Drosophila
44	43.5	49.4	399	6	ABW70654 Recombina
45	43	48.9	569	4	AAW79339 Human pro

ALIGNMENTS

RESULT 1
ID AAW61515
AAW61515 standard; peptide; 19 AA.
AC AAW61515;
XX 26-OCT-1998 (first entry)
DE Peptide cFA, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
KW solid phase synthesis; peptide amide; polyclonal antibody;
KW monoclonal antibody.
XX Synthetic.
OS Homo sapiens.
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WEIENSCHAPPEN.
XX Van Venrooij JMW, Schellekens GA, Raats JMH, Hoet RMA;
MPI, 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX Disclosure, Page 6; 19pp; English.
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX Sequence 19 AA;
XX SQ

Query Match 96.6%; Score 85; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 3.9e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 2
 AAW61517
 ID AAW61517 standard; peptide; 19 AA.
 XX
 AC AAW61517;

DT 26-OCT-1998 (first entry)
 XX
 DE Peptide cfo, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX W09822503-A2.

PD 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.7e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 3
 AAW61509
 ID AAW61509 standard; peptide; 19 AA.
 XX
 AC AAW61509;

XX 26-OCT-1998 (first entry)
 DT
 XX Peptide cfc5, based on cDNA of a profilaggrin repeat.

DE Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 XX solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX W09822503-A2.

PD 28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.

XX 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.

PT Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

XX PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX SQ Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 94.7%; Pred. No. 5.7e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOESTGRSGRSGSGS 19
 |||||
 DB 1 SHOESTGRSGRSGSGS 19

RESULT 4
 AAW61513
 ID AAW61513 standard; peptide; 19 AA.

XX AAW61513;

DT 26-OCT-1998 (first entry)

XX Peptide cfc9, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Modified-site 7 /note= "Citruiline"
FT Modified-site 16 /note= "Citruiline"
FT Modified-site 16 /note= "Citruiline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX Sequences AAM61505-M61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-07;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTXGRSGRSGSGS 19
XX 1 SHOESTXGRSGRSGSGS 19
XX DB
XX
XX RESULT 5
XX AAM61505
XX ID AAM61505 standard; peptide; 19 AA.
XX AC AAM61505;
XX 26-OCT-1998 (first entry)
XX Peptide cft1, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX Modified-site 7 /note= "Citruiline"
XX WO9822503-A2.
XX 28-MAY-1998.
XX

```

```

PF 14-NOV-1997; 97WO-NL000624.
XX 15-NOV-1996; 96NL-01004539.
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX Disclosure; Page 6; 19pp; English.
XX Sequences AAM61505-M61520 are peptides derived from the C-terminal end of
XX the profilaggrin antigen which is recognised by autoantibodies from
XX patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX RA patient's autoimmune antibodies which are reactive with profilaggrin.
XX The peptides were created by using standard solid phase synthesis, which
XX produced them as peptide amides. These sequences may be used in the
XX detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
XX
XX Query Match 95.5%; Score 84; DB 2; Length 19;
XX Best Local Similarity 94.7%; Pred. No. 5.7e-07;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 SHOESTXGRSGRSGSGS 19
XX 1 SHOESTXGRSGRSGSGS 19
XX DB
XX
XX RESULT 6
XX AAM61514
XX ID AAM61514 standard; peptide; 19 AA.
XX AC AAM61514;
XX 26-OCT-1998 (first entry)
XX Peptide cf, based on cDNA of a profilaggrin repeat.
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX Modified-site 7 /note= "Citruiline"
XX WPI; 1998-398613/34.
XX Peptide derived from an antigen recognised by autoantibodies - is
XX reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX used in diagnosis of the disease.
XX

```

PS Disclosure; Page 6; 19pp; English.

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

CC Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5,7e-07;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYGRSRGRSGSGS 19
 DB 1 SHQESTYGRSRGRSGSGS 19

RESULT 7

ID AAW61516 standard; peptide; 19 AA.

AC AAW61516;

DT 26-OCT-1998 (first entry)

DE Peptide cfe, based on cDNA of a profilaggrin repeat.

KA Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

KM monoclonal antibody.

OS Synthetic.

OS Homo sapiens.

PN W09822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97MO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PI (TEWE-) STICHTING TECH WETENSCHAPEN.

PI Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

CC Sequence 19 AA;

Query Match 95.5%; Score 84; DB 2; Length 19;

Best Local Similarity 89.5%; Pred. No. 5,7e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYGRSRGRSGSGS 19
 DB 1 SHQESTYGRSRGRSGSGS 19

RESULT 8

ID ABB97605 standard; protein; 1467 AA.

AC ABB97605;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

KA Human; anti-neoplastic; vulnary; anti-inflammatory; immunomodulator;

KW anti-infective; cerebroprotective; cytostatic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; EST;

OS expressed sequence tag.

OS Homo sapiens.

PN W0200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABB97605.

PT An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 873; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

CC Sequence 1467 AA;

Query Match 90.9%; Score 80; DB 5; Length 1467;

Best Local Similarity 84.2%; Pred. No. 0.0002;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQESTYGRSRGRSGSGS 19
 DB 773 SHQESTYGRSRGRSGSGS 791

RESULT 9

ID AAW61506 standard; peptide; 19 AA.

AC AAW61506;

DT 27-JUN-2002 (first entry)

DE Novel human protein SEQ ID NO: 873.

KA Human; anti-neoplastic; vulnary; anti-inflammatory; immunomodulator;

KW anti-infective; cerebroprotective; cytostatic; rheumatic; gene therapy;

KM neuroprotective; antiparkinsonian; protein therapy; EST;

OS expressed sequence tag.

OS Homo sapiens.

PN W0200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001MO-US026015.

PR 11-SEP-2000; 2000US-00659671.

PA (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-292408/33.

DR N-PSDB; ABB97605.

PT An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 873; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

CC Sequence 1467 AA;

Query Match 90.9%; Score 80; DB 5; Length 1467;

Best Local Similarity 84.2%; Pred. No. 0.0002;

Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DT	26-OCT-1998	(first entry)	
DE	Peptide cfc2, based on cDNA of a profilaggrin repeat.		
KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;		
KM	solid phase synthesis; peptide amide; polyclonal antibody;		
KX	monoclonal antibody.		
OS	Synthetic.		
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
XX	Modified-site	9	
XX		/note="Citruilina"	
XX	MO9822503-AA2.		
XX	28-MAY-1998.		
XX	14-NOV-1997;	97MO-NL000624.	
XX	15-NOV-1996;	96NL-01004539.	
XX	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.		
XX	(TEWE-) STICHTING TECH WETENSCHAPPEN.		
XX	Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA;		
XX	WPI; 1998-398613/34.		
XX	Peptide derived from an antigen recognised by autoantibodies - is		
XX	reactive with autoimmune antibodies from rheumatoid arthritis, and may be		
XX	used in diagnosis of the disease.		
XX	Disclosure; Page 6; 19pp; English.		
XX			
XX	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of		
XX	the profilaggrin antigen which is recognised by autoantibodies from		
XX	patients with rheumatoid arthritis (RA). This peptide is reactive with a		
XX	RA patient's autoimmune antibodies which are reactive with profilaggrin.		
XX	The peptides were created by using standard solid phase synthesis, which		
XX	produced them as peptide amides. These sequences may be used in the		
XX	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as		
XX	for obtaining polyclonal and monoclonal antibodies		
XX			
XX	Sequence 19 AA;		
XX			
XX	Query Match	88.6%; Score 78; DB 2; Length 19;	
XX	Best Local Similarity	84.2%; Pred. No. 5.6e-06;	
XX	Matches	16; Conservative 0; Mismatches 3; Indels 0; Gaps 0	
XX			
XX	QY	1 SHQESTYGRSGRSGSGS 19	
XX			
XX			
XX			
XX			
XX			
XX	DB	1 SHQESTGRSGRSGRSGS 19	
XX			
XX	RESULT 10		
XX	AAW61508		
XX	ID	AAW61508 standard; peptide; 19 AA.	
XX	AC	AAW61508;	
XX	DT	26-OCT-1998 (first entry)	
XX	DE	Peptide cfc4, based on cDNA of a profilaggrin repeat.	
XX	KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
XX	KM	solid phase synthesis; peptide amide; polyclonal antibody;	
XX	KX	monoclonal antibody.	
XX	OS	Synthetic.	
XX	OS	Homo sapiens.	
XX			

XX	Key	Location/Qualifiers
FT	Modified-site	13
FT		/note= "Citrulline"
XX		
XX	W09822503-A2.	
XX		
XX	28-MAY-1998.	
XX		
XX	14-NOV-1997;	97WO-NL000624.
XX		
XX	15-NOV-1996;	96NL-01004539.
XX		
XX	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.	
XX	(TEWE-) STICHTING TECH WETENSCHAPPEN.	
XX		
XX	Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;	
XX		
XX	WPI; 1998-398613/34.	
XX		
XX	Peptide derived from an antigen recognised by autoantibodies - is	
XX	reactive with autoimmune antibodies from rheumatoid arthritis, and may be	
XX	used in diagnosis of the disease.	
XX		
XX	Disclosure; Page 6; 19pp; English.	
XX		
XX	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of	
XX	the profilaggrin antigen which is recognised by autoantibodies from	
XX	patients with rheumatoid arthritis (RA). This peptide is reactive with a	
XX	RA patient's autoimmune antibodies which are reactive with profilaggrin.	
XX	The peptides were created by using standard solid phase synthesis, which	
XX	produced them as peptide amides. These sequences may be used in the	
XX	detection of RA autoimmune antibodies, in the diagnosis of RA, as well as	
XX	for obtaining polyclonal and monoclonal antibodies	
XX		
XX	Sequence 19 AA;	
XX		
XX		
XX	Query Match	88.6%; Score 78; DB 2; Length 19;
XX	Best Local Similarity	84.2%; Pred. No. 5.6e-06;
XX	Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0.	
XX		
XX	1 SHOESTXGRSRGSGSGS 19	
XX	1 SHOESTRGRSRGSGSGS 19	
XX		
XX	RESULT 11	
XX	AAW61507	
XX	ID AAW61507 standard; peptide; 19 AA.	
XX	AAW61507;	
XX		
XX	26-OCT-1998 (first entry)	
XX		
XX	Peptide cfc3, based on cDNA of a profilaggrin repeat.	
XX		
XX	Antigen: autoantibody; rheumatoid arthritis; RA; autoimmune antibody;	
XX	solid phase synthesis; peptide amide; polyclonal antibody;	
XX	monoclonal antibody.	
XX		
XX	Synthetic.	
XX	OS Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
XX	Modified-site	11
XX		/note= "Citrulline"
XX		
XX	W09822503-A2.	
XX		
XX	28-MAY-1998.	
XX		
XX	14-NOV-1997;	97WO-NL000624.
XX		
XX	15-NOV-1996;	96NL-01004539.
XX		

```

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
CC
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 84.2%; Pred. No. 5.6e-06;
XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SHOESTXGRSRGRSGSGS 19
DB 1 SHOESTXGRSRGRSGSGS 19
XX
XX RESULT 12
XX AAW61512
XX ID AAW61512 standard; peptide; 19 AA.
XX
XX AAW61512;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc8, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7 /note= "Citruilline"
XX Modified-site 13 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 28-MAY-1998.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT

```

```

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC The peptides were created by using standard solid phase synthesis, which
CC produced them as peptide amides. These sequences may be used in the
CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
CC for obtaining polyclonal and monoclonal antibodies
CC
XX
XX Sequence 19 AA;
SQ
XX
XX Query Match 88.6%; Score 78; DB 2; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 5.6e-06;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHOESTXGRSRGRSGSGS 19
DB 1 SHOESTXGRSRGRSGSGS 19
XX
XX RESULT 13
XX AAW61511
XX ID AAW61511 standard; peptide; 19 AA.
XX
XX AAW61511;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfc7, based on cDNA of a profilaggrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7 /note= "Citruilline"
XX Modified-site 11 /note= "Citruilline"
XX
XX WO9822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI, 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
CC the profilaggrin antigen which is recognised by autoantibodies from
CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
CC

```

CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.6e-06;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGRSGSGS 19
 |||||
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 14

AAW61510
 ID AAW61510 standard; peptide; 19 AA.

XX AAW61510;

DT 26-OCT-1998 (first entry)

DE Peptide cfc6, based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 7 /note= "Citruiline"

FT Modified-site 9 /note= "Citruiline"

FT WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Page 6; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 19 AA;

Query Match 88.6%; Score 78; DB 2; Length 19;
 Best Local Similarity 89.5%; Pred. No. 5.6e-06;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSRGRSGSGS 19
 |||||
 DB 1 SHQSTXGRSRGRSGSGS 19

RESULT 15
 AAW61520
 ID AAW61520 standard; peptide; 21 AA.

XX AAW61520;

DT 26-OCT-1998 (first entry)

DE Peptide XI based on cDNA of a proflaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KW solid phase synthesis; peptide amide; polyclonal antibody;
 KM monoclonal antibody.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 9 /note= "Citruiline"

FT WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 (TEWE-) STICHTING TECH WETENSCHAPPEN.

PI Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA;

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PT used in diagnosis of the disease.

PS Disclosure; Fig 1; 19pp; English.

XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the proflaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with proflaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies

XX Sequence 21 AA;

Query Match 85.2%; Score 75; DB 2; Length 21;
 Best Local Similarity 88.9%; Pred. No. 1.9e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HQESTXGRSRGRSGSGS 19
 |||||
 DB 4 HQESTXGRSRGRSGSGS 21

Search completed: September 28, 2004, 06:24:20
 Job time : 56.1302 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 ; Search time 111.526 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-9
Perfect score: 88
Sequence: 1 SHQESTXGRSRGRSGXSGS 19

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1349238 segs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	85	96.6	19 9 US-09-308-150-12	Sequence 12, Appl
2	84	95.5	19 9 US-09-308-150-1	Sequence 1, Appl
3	84	95.5	19 9 US-09-308-150-5	Sequence 5, Appl
4	84	95.5	19 9 US-09-308-150-9	Sequence 9, Appl
5	84	95.5	19 9 US-09-308-150-11	Sequence 11, Appl
6	84	95.5	19 9 US-09-308-150-13	Sequence 13, Appl
7	84	95.5	19 9 US-09-308-150-14	Sequence 14, Appl
8	84	95.5	19 9 US-09-308-150-2	Sequence 2, Appl
9	84	95.5	19 9 US-09-308-150-3	Sequence 3, Appl
10	84	95.5	19 9 US-09-308-150-4	Sequence 4, Appl
11	84	95.5	19 9 US-09-308-150-6	Sequence 6, Appl
12	84	95.5	19 9 US-09-308-150-7	Sequence 7, Appl
13	84	95.5	19 9 US-09-308-150-8	Sequence 8, Appl
14	84	95.5	21 9 US-09-308-150-10	Sequence 10, Appl
15	75	85.2	22 9 US-09-747-029A-22	Sequence 22, Appl

16	50	56.8	477	15	US-10-161-927-62	Sequence 62, Appl
17	47	53.4	123	16	US-10-767-701-32436	Sequence 32436, A
18	47	53.4	436	14	US-10-156-761-13022	Sequence 13022, A
19	47	53.4	838	16	US-10-437-963-146503	Sequence 146503, A
20	47	53.4	854	12	US-10-425-114-57838	Sequence 57838, A
21	46	52.3	78	16	US-10-437-963-203073	Sequence 203073, A
22	45	51.1	653	9	US-09-746-801A-13	Sequence 13, Appl
23	44	50.0	262	12	US-10-425-114-72166	Sequence 72166, A
24	44	50.0	262	12	US-10-425-114-72167	Sequence 72167, A
25	44	50.0	563	16	US-10-437-963-198539	Sequence 198539, A
26	44	50.0	691	9	US-09-919-585-6	Sequence 6, Appl
27	44	50.0	722	15	US-10-274-184-4	Sequence 4, Appl
28	44	50.0	722	16	US-10-760-407-4	Sequence 4, Appl
29	44	50.0	724	9	US-09-919-585-9	Sequence 9, Appl
30	44	50.0	724	15	US-10-274-194-2	Sequence 2, Appl
31	44	50.0	724	16	US-10-760-407-2	Sequence 2, Appl
32	44	50.0	745	12	US-10-260-708-79	Sequence 79, Appl
33	44	50.0	745	14	US-10-195-101-16	Sequence 16, Appl
34	44	50.0	745	14	US-10-161-565-24	Sequence 24, Appl
35	44	50.0	936	14	US-10-156-761-11212	Sequence 11212, A
36	44	50.0	1087	9	US-09-918-909-24	Sequence 24, Appl
37	44	50.0	1087	16	US-10-641-991-24	Sequence 24, Appl
38	43	48.9	91	16	US-10-437-963-132029	Sequence 132029, A
39	43	48.9	130	16	US-10-437-963-174101	Sequence 174101, A
40	43	48.9	134	12	US-10-425-114-43265	Sequence 43265, A
41	43	48.9	134	12	US-10-425-114-63280	Sequence 63280, A
42	43	48.9	182	12	US-10-424-599-61480	Sequence 21480, A
43	43	48.9	500	16	US-10-437-963-190244	Sequence 190244, A
44	43	48.9	570	10	US-09-847-102A-43	Sequence 43, Appl
45	43	48.9	611	16	US-10-437-963-149539	Sequence 149539, A

ALIGNMENTS

RESULT 1
US-09-308-150-12
Sequence 12, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoel, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY ANTIBODY BODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS, TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Derived from US-09-308-150-12
OTHER INFORMATION: Known CDNA sequences of human profilaggrin

Query Match 96.6%; Score 85; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 5.7e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 SHQESTXGRSRGRSGXSGS 19

Db 1 SHQSTXGRSGRSGSGS 19

RESULT 2

US-09-308-150-1
Sequence 1, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoest, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 3

US-09-308-150-5
Sequence 5, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoest, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 8.1e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 4

US-09-308-150-9
Sequence 9, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoest, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match 95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 SHQSTXGRSGRSGSGS 19
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 5

US-09-308-150-11
Sequence 11, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoest, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 19

```

; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-11

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8,1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTXGRSRGRSGSGS 19

```

```

RESULT 6
US-09-308-150-13
; Sequence 13, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-13

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8,1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTXGRSRGRSGSGS 19

```

```

RESULT 7
US-09-308-150-14

```

```

; Sequence 14, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 14
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-14

```

```

Query Match      95.5%; Score 84; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 8,1e-06;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 SHOESTXGRSRGRSGSGS 19
Db      1 SHOESTXGRSRGRSGSGS 19

```

```

RESULT 8
US-09-308-150-2
; Sequence 2, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Maltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: Known cDNA sequences of human profilaggrin
US-09-308-150-2

```

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 6,4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
|||
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 9
US-09-308-150-3
Sequence 3, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 6,4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
|||
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 10
US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 6,4e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
|||
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 11
US-09-308-150-6
Sequence 6, Application US/09308150

Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoet, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
APPLICANT: Stichting voor de Technische Wetenschappen
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT APPLICATION NUMBER: US/09/308,150
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match 88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6,4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQSTXGRSGRSGSGS 19
|||
Db 1 SHQSTXGRSGRSGSGS 19

RESULT 12
US-09-308-150-7
Sequence 7, Application US/09308150
Patent No. US20020137092A1
GENERAL INFORMATION:
APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius

```

; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Cy      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTXGRSGRSGSGS 19

```

```

RESULT 13
US-09-308-150-8
; Sequence 8, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

```

```

Query Match      88.6%; Score 78; DB 9; Length 19;
Best Local Similarity 89.5%; Pred. No. 6.4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Cy      1  SHOESTXGRSGRSGSGS 19
Db      1  SHOESTXGRSGRSGSGS 19

```

```

RESULT 14
US-09-308-150-10
; Sequence 10, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Ventrcooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; APPLICANT: Stichting voor de Technische Wetenschappen
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM
; FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/NL97/00624
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
; NAME/KEY: DISULFID
; LOCATION: (3)..(16)
US-09-308-150-10

```

```

Query Match      85.2%; Score 75; DB 9; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.0002;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Cy      2  HOESTXGRSGRSGSGS 19
Db      4  HOESTXGRSGRSGSGS 21

```

```

RESULT 15
US-09-747-029A-22
; Sequence 22, Application US/09747029A
; Patent No. US20020143143A1
; GENERAL INFORMATION:
; APPLICANT: Union, Ann
; APPLICANT: Moereels, Henri
; APPLICANT: Meheus, Lydie
; TITLE OF INVENTION: PEPTIDES DESIGNED FOR THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 11362.0031NPU500 INNS:031
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: EP 00870195.5
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: EP 99870280.7
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 22
; TYPE: PRT

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 14.4479 Seconds
(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-9
Perfect score: 88
Sequence: 1 SHOESTXGRSGRSGXSGS 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	51.1	653	US-09-513-057C-13	Sequence 13, Appl
2	44	50.0	722	US-09-984-890-4	Sequence 4, Appl
3	44	50.0	724	US-09-984-890-2	Sequence 2, Appl
4	44	50.0	745	US-09-523-849-36	Sequence 36, Appl
5	42	47.7	169	US-09-342-084-6	Sequence 6, Appl
6	42	47.7	448	US-09-252-991A-24066	Sequence 24066, A
7	42	47.7	722	US-08-817-832B-32	Sequence 32, Appl
8	42	47.7	2237	US-08-354-973-1	Sequence 1, Appl
9	41	46.6	255	US-09-252-991A-19598	Sequence 19598, A
10	41	46.6	294	US-09-252-991A-21947	Sequence 21947, A
11	41	46.6	326	US-09-252-991A-20075	Sequence 20075, A
12	41	46.6	363	US-09-252-991A-25132	Sequence 25132, A
13	41	46.6	409	US-09-489-039A-14217	Sequence 14217, A
14	41	46.6	421	US-09-252-991A-3236	Sequence 3236, A
15	41	46.6	480	US-09-252-991A-22879	Sequence 22879, A
16	41	46.6	629	US-09-252-991A-22901	Sequence 22901, A
17	40.5	46.0	258	US-10-164-595-55	Sequence 59, Appl
18	40.5	46.0	274	US-09-976-594-417	Sequence 417, Appl
19	40.5	46.0	378	US-10-164-595-2	Sequence 2, Appl
20	40	45.5	239	US-09-252-991A-28985	Sequence 28985, A
21	40	45.5	255	US-09-342-084-11	Sequence 11, Appl
22	40	45.5	468	US-09-252-991A-19026	Sequence 19026, A
23	40	45.5	614	US-09-252-991A-29695	Sequence 29695, A
24	39.5	44.9	257	US-09-252-991A-22204	Sequence 22204, A
25	39.5	44.9	371	US-09-252-991A-29985	Sequence 29985, A
26	39	44.3	189	US-09-252-991A-24576	Sequence 24576, A
27	39	44.3	195	US-09-252-991A-30082	Sequence 30082, A

28	39	44.3	288	US-09-252-991A-27979	Sequence 27979, A
29	39	44.3	326	US-09-252-991A-17002	Sequence 17002, A
30	39	44.3	486	US-08-821-355A-8	Sequence 8, Appl
31	39	44.3	486	US-09-003-687A-8	Sequence 8, Appl
32	39	44.3	486	US-09-136-605-8	Sequence 8, Appl
33	39	44.3	491	US-09-489-039A-7836	Sequence 7836, Appl
34	38	43.2	154	US-09-252-991A-21434	Sequence 21434, A
35	38	43.2	157	US-09-252-991A-27836	Sequence 27836, A
36	38	43.2	274	US-09-252-991A-16696	Sequence 16696, A
37	38	43.2	279	US-09-252-991A-32606	Sequence 32606, A
38	38	43.2	329	US-09-252-991A-28088	Sequence 28088, A
39	38	43.2	379	US-09-252-991A-28428	Sequence 28428, A
40	38	43.2	380	US-09-252-991A-24774	Sequence 24774, A
41	38	43.2	392	US-09-252-991A-31291	Sequence 31291, A
42	38	43.2	447	US-09-189-637A-351	Sequence 351, Appl
43	38	43.2	508	US-08-818-024-3	Sequence 3, Appl
44	38	43.2	508	US-09-314-775A-3	Sequence 3, Appl
45	38	43.2	518	US-09-252-991A-25967	Sequence 25967, A

ALIGNMENTS

RESULT 1
US-09-513-057C-13
Sequence 13, Application US/09513057C
Patent No. 6433251
GENERAL INFORMATION:
APPLICANT: Wagner, et al.
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM
FILE REFERENCE: 1505-54357
CURRENT APPLICATION NUMBER: US/09/513,057C
CURRENT FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 653
TYPE: PRT
ORGANISM: Cardamine hirsputa
US-09-513-057C-13

Query Match 51.1%; Score 45; DB 4; Length 653;
Best Local Similarity 52.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

CY 1 SHOESTXGRSGRSGXSGS 19
DB 573 SROVSTSSASGREGISGS 591
RESULT 2
US-09-984-890-4
Sequence 4, Application US/09984890
Patent No. 6492156
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CLO01306
CURRENT APPLICATION NUMBER: US/09/984,890
CURRENT FILING DATE: 2001-10-31
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 722
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match 50.0%; Score 44; DB 4; Length 722;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
Db 562 SPGSHSGRGRGASGS 576

RESULT 3
US-09-984-890-2
; Sequence 2, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-890-2

Query Match 50.0%; Score 44; DB 4; Length 724;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
Db 564 SPGSHSGRGRGASGS 578

RESULT 4
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scaccheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match 50.0%; Score 44; DB 4; Length 745;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 STXGRGRGRGXSGS 19
| | | | |
Db 585 SPGSHSGRGRGASGS 599

RESULT 5
US-09-342-084-6

; Sequence 6, Application US/09342084
; Patent No. 6251668
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: Transcription Coactivators
; FILE REFERENCE: BB-1169-A
; CURRENT APPLICATION NUMBER: US/09/342,084
; CURRENT FILING DATE: 1999-06-28
; EARLIER APPLICATION NUMBER: 60/092,659
; EARLIER FILING DATE: July 13, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (128)
US-09-342-084-6

Query Match 47.7%; Score 42; DB 3; Length 169;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 SHQSTYGRGRGRGXSGS 19
| | | | |
Db 15 NKKSSGSSRGRTSPGS 33

RESULT 6
US-09-252-991A-24066
; Sequence 24066, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24066
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24066

Query Match 47.7%; Score 42; DB 4; Length 448;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 SHQSTYGRGRGRGXSG 18
| | | | |
Db 412 SHQGPAGRPGRPRPPG 429

RESULT 7
US-08-817-832B-32
; Sequence 32, Application US/08817832B
; Patent No. 6579691
; GENERAL INFORMATION:
; APPLICANT: MANDELKOW, Eckhard, et al.
; TITLE OF INVENTION: NO. 6579691a1 Protein Kinase (NPK-110)
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower

DB 22 SKSPSRGRSGPAG 35

RESULT 11

US-09-252-991A-20075
; Sequence 20075, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20075
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (15)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20075

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 326;
Best Local Similarity 64.3%; Pred. No. 42;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TXGRSRGRSGXSGS 19

DB 272 TSPSRGRSGXSGS 285

RESULT 12

US-09-252-991A-25132
; Sequence 25132, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25132
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25132

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 363;
Best Local Similarity 53.3%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 SHOESTYGRSRGRSGXSGS 15

DB 91 AHHSVGRGRGRWARG 105

RESULT 13

US-09-489-039A-14217
; Sequence 14217, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14217
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14217

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 409;
Best Local Similarity 47.4%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 SHOESTYGRSRGRSGXSGS 19

DB 374 SROASSPNEBSKRGSESGS 392

RESULT 14

US-09-252-991A-32326
; Sequence 32326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32326
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32326

Query Match

Best Local Similarity 46.6%; Score 41; DB 4; Length 421;
Best Local Similarity 47.1%; Pred. No. 54;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 HOESTYGRSRGRSGXSGS 18

DB 13 HARTGTGTGRGRGRSG 29

RESULT 15

US-09-252-991A-22879
; Sequence 22879, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22879

Blank

Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22; Search time 14.4375 Seconds
(without alignments)
139.915 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HQHOESTXGRSRGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	86	74.8	2248	2 A35938
2	79	68.7	416	2 A32947
3	70	60.9	591	2 A45135
4	53	46.1	822	2 T51049
5	51	44.3	506	1 W2WL47
6	48.5	42.2	524	2 S28539
7	48.5	42.2	670	2 I65967
8	48	41.7	214	2 T45714
9	47	40.9	266	2 T14345
10	47	40.9	266	2 T14348
11	47	40.9	400	2 B48613
12	47	40.9	730	2 S24376
13	46.5	40.4	338	2 B75584
14	46	40.0	136	2 T35632
15	46	40.0	268	2 T14341
16	46	40.0	849	2 A96592
17	45	39.1	202	2 T50635
18	45	39.1	268	2 T14344
19	45	39.1	289	2 AD3394
20	45	39.1	384	1 G70961
21	44	38.3	440	2 T24232
22	44	38.3	825	2 UC4163
23	44	38.3	836	2 G84727
24	44	38.3	1215	2 I52882
25	44	38.3	1804	2 T34518
26	43	37.4	38	2 S74088
27	43	37.4	334	2 G81744
28	43	37.4	1219	2 T61713
29	43	37.4	1669	1 CGWS4B

30	42.5	37.0	299	2	T35765	hypothetical prote
31	42	36.5	123	2	T16234	hypothetical prote
32	42	36.5	323	2	D85074	hypothetical prote
33	42	36.5	424	2	T33552	hypothetical prote
34	42	36.5	519	2	T22091	hypothetical prote
35	42	36.5	682	2	T15092	hypothetical prote
36	42	36.5	888	2	S78288	preprotein translo
37	42	36.5	952	2	T18900	disintegrin and me
38	41.5	36.1	292	2	T03122	hypothetical prote
39	41.5	36.1	487	2	T06040	hypothetical prote
40	41	35.7	229	2	UC7219	nuclear protein SR
41	41	35.7	294	2	A96155	hypothetical prote
42	41	35.7	312	2	A31846	130K paracrysta111
43	41	35.7	350	2	A85056	probable transposo
44	41	35.7	368	2	E87753	protein C43E11.5 (
45	41	35.7	368	2	T29779	hypothetical prote

ALIGNMENTS

RESULT 1
A35938
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence__revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; PMID:9106347; PMID:2248957
A:Accession: A35938
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2248 <GAN>
A:Cross-references: GB:J02929
C:Genetics:
A:Gene: GDB:FLG
A:Cross-references: GDB:119912; OMIM:135940
A:Map position: 1q21-1q21
C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: BP hand; epidermis; polymorphism; tandem repeat
F:245-569/Region: profilaggrin repeat
F:370-893/Region: profilaggrin repeat
F:1074-1397/Region: profilaggrin repeat
F:1573-1896/Region: profilaggrin repeat

Query Match 74.8%; Score 86; DB 2; Length 2248;
Best local similarity 81.0%; Pred. No. 0.00016;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 HQHOESTXGRSRGRCGRSGS 21
Db 225 HOSHQESTXGRSRGRCGRSGS 245

RESULT 2
A32947
profilaggrin precursor - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Dec-1989 #sequence__revision 04-Sep-1992 #text_change 29-Sep-1999
C:Accession: A32947
R:McKinley-Grant, L.J.; Idler, W.W.; Bernstein, I.A.; Parry, D.A.D.; Cammizzaro, L.; Croc
Proc. Natl. Acad. Sci. U.S.A. 86, 4848-4852, 1989
A:Title: Characterization of a cDNA clone encoding human profilaggrin and localization of tr
A:Reference number: A32947; PMID:9296901; PMID:2740331
A:Accession: A32947
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <MCK>
A:Cross-references: GB:M24355; NID:g182604; PIND:AA52454.1; PTD:g182605
A>Note: the authors translated the codon CAC for residue 188 as Gln, and AAT for residue
C:Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat

Query Match 68.7%; Score 79; DB 2; Length 416;
 Best Local Similarity 80.0%; Pred. No. 0.00045;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QCHOESTXGRSGRCGRSGS 21
 Db 6 QSHOESTRGRSAGRSGRSGS 25

RESULT 3
 A45135
 C:Species: Homo sapiens (man)
 C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A45135

R:Presland, R.B.; Haydock, P.V.; Fleckman, P.; Nitsunskisiri, W.; Dale, B.A.
 J. Biol. Chem. 267, 23772-23781, 1992
 A:Title: Characterization of the human epidermal profilaggrin gene. Genomic organization
 A:Reference number: A45135; MUID:93054736; PMID:1429717

A:Accession: A45135
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-591 <PRE>
 A:Cross-references: GB:101089; GB:M90967; NID:9190408; PIDN:AAA60177.1; PID:G553621
 A:Note: sequence extracted from NCBI backbone (NCBIP:118773)
 C:Genetics:

A:Gene: GDB:FLG
 A:Cross-references: GDB:119912; OMIM:135940
 A:Map position: 1q21-1q21
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat
 F:49-81/Domain: calmodulin repeat homology <EP2>

Query Match 60.9%; Score 70; DB 2; Length 591;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 QCHOESTXGRSGRCGRSGS 21
 Db 448 QSHOESTRGRSAGRSGRSGS 467

RESULT 4
 T51049
 related to nucleolar phosphoprotein [imported] - Neurospora crassa
 N:Alternate names: protein B12P1.10
 C:Species: Neurospora crassa
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51049
 R:Schulte, U.; Aign, V.; Hehseisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: 225286
 A:Accession: T51049
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-822 <SCH>
 A:Cross-references: EMBL:AJ390091; GSPDB:GN00116; NCSP:B12P1.10
 C:Experimental source: BAC clone B12P1; strain OR4A
 C:Genetics:
 A:Gene: NCSP:B12P1.10
 A:Map position: 6
 A:Introns: 80/2

Query Match 46.1%; Score 53; DB 2; Length 822;
 Best Local Similarity 58.8%; Pred. No. 4.6;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 QESTXGRSGRCGRSGS 21
 Db 414 RESACRTRGRGRCGCT 430

RESULT 5

W2ML47
 E2 protein - human papillomavirus type 47
 C:Species: human papillomavirus type 47
 A:Note: host Homo sapiens (man)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: D35324

R:Kiyono, T.; Adachi, A.; Ichibaashi, M.
 Virology 177, 401-405, 1990
 A:Title: Genome organization and taxonomic position of human papillomavirus type 47 infer
 A:Reference number: A35324; MUID:90281611; PMID:2162112
 C:Accession: D35324

A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-506 <KIT>
 A:Cross-references: GB:M32305; NID:9333062; PIDN:AAA4679.1; PID:9333067
 C:Superfamily: Papillomavirus E2 protein
 C:Keywords: DNA binding; early protein; transcription regulation

Query Match 44.3%; Score 51; DB 1; Length 506;
 Best Local Similarity 66.7%; Pred. No. 6.1;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 7 STXGRSGRCGRSGS 21
 Db 346 NTRGRGRGRCGRSAGS 360

RESULT 6
 S38539
 disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 1 - human
 N:Alternate names: ADAM1; MDC
 C:Species: Homo sapiens (man)
 C:Date: 07-Oct-1994 #sequence_revision 17-Nov-1995 #text_change 26-May-2000
 C:Accession: S38539; I52965

R:Emi, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, J.; Ito, I.; Kasumi, F.; Nakamui
 Nature Genet. 5, 151-157, 1993
 A:Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearrang
 A:Reference number: S38539; MUID:94073190; PMID:8252040
 C:Accession: S38539
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-524 <EMI>
 A:Cross-references: GB:D17390; NID:9452188; PIDN:BA404213.1; PID:d1004732; PID:g484255

R:Katagiri, T.; Harada, Y.; Emi, M.; Nakamura, Y.
 Cytogenet. Cell Genet. 68, 39-44, 1995
 A:Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and
 A:Reference number: I52965; MUID:95044425; PMID:7956356
 C:Accession: I52965

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 7-225, 'D', 227-524 <KAT>
 A:Cross-references: GB:D11872; NID:9505090; PIDN:BA406671.1; PID:d1007243; PID:g836684
 C:Comment: For an alternative splice form, see PIR:I65967.
 C:Genetics:

A:Gene: GDB:MDC; ADAM1
 A:Cross-references: GDB:230267; OMIM:155120
 A:Map position: 17q21.3-17q21.3
 A:Introns: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341/
 C:Superfamily: disintegrin homology
 C:Keywords: alternative splicing; hydrolase; metalloproteinase
 F:344-427/Domain: disintegrin homology <DIS>

Query Match 42.3%; Score 48.5; DB 2; Length 524;
 Best Local Similarity 45.0%; Pred. No. 14;
 Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHOESTXGRSGRCGRSGS 21

Db 468 CYEKLNVETGERSGCRKGS 487

RESULT 7

165967
disintegrin-like metalloproteinase (EC 3.4.24.-), splice form 2 - human
N/Alternate names: ADAM11; MDC
C/Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 17-Mar-2000 #text_change 11-Jan-2002
C/Accession: 165967; S38539
R/Katagiri, T.; Harada, Y.; Eml, M.; Nakamura, Y.
Cytogenet. Cell Genet. 68, 39-44, 1995
A/Title: Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron organization and
A/Reference number: 152965; MUID:95044425; PMID:7956356
A/Accession: 165967
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 7-670 <KAT>
A/Cross-references: GB:DJ3872; NID:9505090; PIDN:BA06670.1; PID:9836683
R/Eml, M.; Katagiri, T.; Harada, Y.; Saito, H.; Inazawa, Y.; Ito, I.; Kasumi, F.; Nakamu
Nature Genet. 5, 151-157, 1993
A/Title: A novel metalloproteinase/disintegrin-like gene at 17q21.3 is somatically rearran
A/Reference number: S38539; MUID:94073190; PMID:8252040
A/Accession: S38539
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-228; N., 227-495, 'PQGRPAWLPDLCQLHSSSARPGGRHQ' <EMT>
A/Cross-references: GB:DJ7390; NID:9452188; PIDN:BA04213.1; PID:9484255
C/Comment: For an alternative splice form, see PIR:S38539.
C/Genetics:
A/Gene: GDB:MDC; ADAM11
A/Cross-references: GDB:230267; OMIM:155120
A/Map position: 17q21.3-17q21.3
A/Introns: 28/3; 57/2; 82/3; 105/1; 127/3; 152/3; 176/3; 232/2; 260/3; 291/1; 308/2; 341
C/Superfamily: disintegrin homology
C/Keywords: alternative splicing; hydrolase; metalloproteinase
F.344-427/Domain: disintegrin homology <Dis>

Query Match 42.2%; Score 48.5; DB 2; Length 670;

Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHQE-STXGRSRGCRSGS 21

Db 468 CYEKLNVETGERSGCRKGS 487

RESULT 8

165967

endochitinase-like protein - Arabidopsis thaliana

N/Alternate names: protein P1P2.90

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000

C/Accession: T45714
R/Choinsse, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999

A/Reference number: 223010

A/Accession: T45714

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-214 <CHO>

A/Cross-references: EMBL:AL12955

A/Experimental source: cultivar Columbia; BAC clone F1P2

C/Genetics:

A/Map position: 3

A/Introns: 93/1

A/Note: P1P2.90
C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

Query Match 41.7%; Score 48; DB 2; Length 214;

Best Local Similarity 56.2%; Pred. No. 8.4;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 HQCHESTXGRSRGRC 16
Db 93 HFCYIEIKKRSRGRC 108

RESULT 9

165967
chitinase (EC 3.2.1.14) EP3-3/E7, class IV - carrot
C/Species: Daucus carota (carrot)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C/Accession: T14345
R/Kragh, K.; De Vries, S.C.
submitted to the EMBL Data Library, March 1996
A/Reference number: 217995
A/Accession: T14345
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-266 <KRA>
A/Cross-references: EMBL:U52847; NID:91549332; PID:91549333
A/Experimental source: strain sg766 trophy
C/Genetics:
A/Gene: EP3
C/Function:
A/Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami
A/Pathway: polysaccharide degradation
C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C/Keywords: glycosidase; hydrolase; polysaccharide degradation
F.21-56/Domain: hevein chitin-binding domain homology <HCB>
F.69-266/Domain: plant chitinase homology <PCH>

Query Match 40.3%; Score 47; DB 2; Length 266;

Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CHQE-STXGRSRGRC 16

Db 136 CHKETNGRDKSYC 149

RESULT 10

165967

probable chitinase (EC 3.2.1.14) EP3B/E6, class IV - carrot

C/Species: Daucus carota (carrot)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000

C/Accession: T14348

R/Kragh, K.; De Vries, S.C.
submitted to the EMBL Data Library, March 1996

A/Reference number: 217995

A/Accession: T14348

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-266 <KRA>

A/Cross-references: EMBL:U52848; NID:91549334; PID:91549335

A/Experimental source: strain sg766 trophy

C/Genetics:

A/Note: EP3

C/Function:

A/Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami
A/Pathway: polysaccharide degradation

C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C/Keywords: glycosidase; hydrolase; polysaccharide degradation

F.21-56/Domain: hevein chitin-binding domain homology <HCB>
F.69-266/Domain: plant chitinase homology <PCH>

Query Match 40.3%; Score 47; DB 2; Length 266;

Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CHQE-STXGRSRGRC 16

Db 136 CHKETNGRDKSYC 149

```

RESULT 11
B48613
surface protein SU - avian myeloblastosis virus
C:Species: avian myeloblastosis virus
C:Date: 21-Jan-1994 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C:Accession: B48613
C:Joliet, V.; Boroughs, K.; Lasserre, F.; Crochet, J.; Dambrine, G.; Smith, R.E.; Perbal
Virology 195, 812-819, 1993
A:Title: Pathogenic potential of myeloblastosis-associated virus: implication of env pro
A:Reference number: A48613; MUID:93311743; PMID:8393249
A:Contents: MAV1(N)/2
A:Accession: B48613
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-400 <UD>
A>Note: sequence extracted from NCBI backbone (NCBI:135488)
C:Superfamily: type C retrovirus env polypeptide

Query Match 40.9%; Score 47; DB 2; Length 400;
Best Local Similarity 38.1%; Pred. No. 19;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
DB 251 YNCSQEGCGCGCQYRCCKARS 271

RESULT 12
S24376
2-aminobenzoyl-CoA monooxygenase/reductase (EC 1.-.-.-) - Pseudomonas sp. plasmid pKB740
C:Species: Pseudomonas sp.
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 08-Oct-1999
C:Accession: S24376; S13857
R:Altenschmidt, U.; Bokranz, M.; Fuchs, G.
Eur. J. Biochem. 207, 715-722, 1992
A:Title: Novel aerobic 2-aminobenzoate metabolism. Nucleotide sequence of the plasmid ca
P.
A:Reference number: S24374; MUID:92339462; PMID:1633822
A:Accession: S24376
A:Molecule type: DNA
A:Residues: 1-730 <ALT1>
A:Cross-references: EMBL:X66604; NID:G45865; PIDD:CAA47170.1; PIDD:G45868
A>Note: the authors translated the codon TGG for residue 680 as Asp, CGT for residue 703
R:Altenschmidt, U.; Eckerskorn, C.; Fuchs, G.
Eur. J. Biochem. 194, 647-653, 1990
A:Title: Evidence that enzymes of a novel aerobic 2-amino-benzoate metabolism in dentrit
A:Reference number: S13857; MUID:91099342; PMID:2176602
A:Accession: S13857
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4, 'X', 6-20, 'D', 21-24 <ALT2>
C:Genetics:
A:Genome: Plasmid
C:Keywords: FAD; flavoprotein; homodimer; NAD; oxidoreductase
F:8-12/Region: FAD binding #status predicted
F:157-162/Region: NAD binding #status predicted

Query Match 40.9%; Score 47; DB 2; Length 730;
Best Local Similarity 52.6%; Pred. No. 31;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 19
DB 332 HQTAAESNIAWRCGRGA 350

RESULT 13
B75584
cytochrome-c peroxidase (EC 1.1.1.5) DRA0301 [similarity] - Deinococcus radiodurans (st
N:Alternate names: methylamine utilization protein
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Apr-2000

```

```

C:Accession: B75584
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathavan, J.; Lam, P.; McDonald, D.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75584
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <WH1>
A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDD:AAF12472.1; PIDD:9646076;
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0301
A:Map position: 2
C:Superfamily: Pseudomonas cytochrome-c peroxidase; Pseudomonas cytochrome-c peroxidase
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:72, 75/Binding site: heme (Cys) (covalent) #status predicted
F:76, 273/Binding site: heme iron (His) (axial ligands) #status predicted
F:216, 219/Binding site: heme (Cys) (covalent) (high potential) #status predicted
F:220/Binding site: heme iron (His) (axial ligand) (high potential) #status predicted

Query Match 40.4%; Score 46.5; DB 2; Length 338;
Best Local Similarity 56.5%; Pred. No. 20;
Matches 13; Conservative 1; Mismatches 4; Indels 5; Gaps 2;

QY 3 CHQES--TXGR--SRGRCGRSG 20
DB 75 CHQSHAFDGRATSLGFRSG 97

RESULT 14
T35632
probable transposase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
C:Accession: T35632
R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21584
A:Accession: T35632
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-136 <SEB>
A:Cross-references: EMBL:AL079356; PIDD:CA845627.1; GSPDB:GN00070; SCQEDB:SC6G9.36C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCQEDB:SC6G9.36C
C:Superfamily: Streptomyces coelicolor probable transposase SC6G9.36C

Query Match 40.0%; Score 46; DB 2; Length 136;
Best Local Similarity 44.4%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 OCHQESTXGRSGRCGRSGS 19
DB 110 RAHQHAAARRGRTGTRA 127

RESULT 15
T14341
probable chitinase (EC 3.2.1.14) EP3-1/H5, class IV - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14341
R:Kraigh, K.; Toonen, M.A.U.; Bucherna, N.; Lo Schiavo, F.; Hendriks, T.; Weijer, E.A.; Ku
Plant Mol. Biol. 31, 631-645, 1996
A:Title: Characterization of carrot chitinases able to rescue the temperature-sensitive
A:Reference number: Z17992; MUID:96382431; PMID:8790295
A:Accession: T14341
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```


A/Residues: 1-268 <KRA>
A/Cross-references: EMBL:U52845; NID:G1549328; PIDN:AAC49435.1; PID:G1549329
A/Experimental source: strain sg766 trophy
C/Genetics:
A/Note: EP3
C/Function:
A/Description: catalyzes the hydrolysis of the beta-1,4-linkages of N-acetyl-D-glucosami
A/Pathway: polysaccharide degradation
C/Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C/Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 40.0%; Score 46; DB 2; Length 268;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 CHOESTXGRSRGRC 16
Db 138 CHOESTSGRDHNYC 151

Search completed: September 28, 2004, 06:15:04
Job time : 16.4375 secs

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 7.98438 Seconds
(without alignments)
136.952 Million cell updates/sec

Title: US-09-308-150-10

Perfect score: 115
Sequence: 1 HQCHQESTXGRSGRGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	79	68.7	F11A_HUMAN	P20330 homo sapien
2	51	44.3	VE2_HPV47	P22420 human papill
3	49.5	43.0	AD11_MOUSE	Q91V44 mus musculu
4	48.5	42.2	AD11_MOUSE	Q75078 homo sapien
5	47.5	41.3	AKP1_RAT	O88884 r a kinase
6	46	40.0	PER2_AZOVI	P82802 azotobacter
7	45	39.1	Z205_HUMAN	O95201 homo sapien
8	44	38.3	5E5_RAT	O63003 rattus norv
9	44	38.3	M2GD_HUMAN	Q9u17 homo sapien
10	43	37.4	DEF1_MYTGA	P80571 mytilus gal
11	43	37.4	CA14_MOUSE	P02463 mus musculu
12	42.5	37.0	AD21_MOUSE	Q91476 mus musculu
13	42	36.5	Y1CG_MOUSE	O8mu7 caenorhabdi
14	42	36.5	LSM4_CAEEL	Q19552 caenorhabdi
15	42	36.5	Y4CG_RHISN	P55389 rhizobium s
16	42	36.5	SECA_ODOSI	P49649 odontella s
17	42	36.5	ARHB_RAT	Q9e667 rattus norv
18	42	36.5	1805_1	Q92445 homo sapien
19	42	36.5	PL15_CHICK	Q98917 gallus gall
20	41	35.7	MRK2_MOUSE	O32522 mycobacteri
21	41	35.7	SEAR_MOUSE	Q50612 mycobacteri
22	41	35.7	808_1	P25764 oryza sativ
23	41	35.7	PHYB_ORYSA	Q02427 drosophila
24	40.5	35.2	RBP1_DROME	Q974K6 thermoplas
25	40.5	35.2	TH1T_THREVO	Q43306 homo sapien
26	40.5	35.2	AD20_HUMAN	Q92766 homo sapien
27	40.5	35.2	RRE1_HUMAN	Q91159 mus musculu
28	40.5	35.2	DEF1_MOUSE	P80154 aescina cyz
29	40	34.8	DEF1_ADSY	Q9h4f8 halobacteri
30	40	34.8	PCRB_HALNI	O8345 treponema p
31	40	34.8	PCNK_TREPA	Q70165 mus musculu
32	40	34.8	FCN1_MOUSE	Q9wec8 rattus norv
33	40	34.8	FCN1_RAT	

34	40	34.8	392	1	NO1C_RHIFR	P26508 rhizobium f
35	40	34.8	562	1	MTRE_MYCLE	Q9ccj1 mycobacteri
36	40	34.8	650	1	BM86_BOOMI	P20736 boophilus m
37	40	34.8	768	1	ITB8_RABIT	P26013 oryctolagus
38	40	34.8	816	1	AD15_RAT	Q9qyV0 r adam 15 p
39	40	34.8	857	1	AKP1_MOUSE	O08715 mus musculu
40	40	34.8	1969	1	Z292_HUMAN	O60281 homo sapien
41	39	33.9	101	1	RS14_VIBCH	Q9kn27 vibrio chol
42	39	33.9	102	1	SAAS_MSAU	P81491 mesocricetu
43	39	33.9	106	1	COLA_HORSE	P02704 equus cabal
44	39	33.9	110	1	PER2_AQUAE	O65511 aquilex aeo
45	39	33.9	224	1	WFD5_HUMAN	O8tcv5 homo sapien

ALIGNMENTS

RESULT 1
FIL1_HUMAN STANDARD; PRT; 416 AA.

AC P20930;1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filaggrin precursor (Fragment).
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McKittrick, Grant L.J., Idler W.W., Bernstein I.A., Parry D.A.D.,
RA Cammizaro L., Croce C.M., Huebner K., Lessin S.R., Steinert P.M.,
RT "Characterization of a cDNA clone encoding human filaggrin and
RT localization of the gene to chromosome region 1q21."
RL Proc. Natl. Acad. Sci. U.S.A. 86:4848-4852(1989).
RM [2]
RN CITRULLINATION.
RP MEDLINE=96374388; PubMed=8780679;
RA Senshu T., Khan S., Ogawa H., Marabe M., Asaga H.;
RT "Preferential delamination of Keratin K1 and filaggrin during the
RT terminal differentiation of human epidermis."
RL Biochem. Biophys. Res. Commun. 225:712-719(1996).
CC -1- FUDCTRON: Aggregates keratin intermediate filaments and promotes
CC disulfide-bond formation among the intermediate filaments during
CC terminal differentiation of mammalian epidermis.
CC -1- PTH: Filaggrin is initially synthesized as a large, insoluble,
CC highly phosphorylated precursor containing many tandem copies of
CC 324 AA, which are not separated by "large linker". The precursor
CC is deposited as keratohyalin granules. During terminal
CC differentiation it is dephosphorylated and proteolytically
CC cleaved.

-1- PTH: Undergoes delamination of some arginine residues
(citrullination).

CC The SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC EMBL; M24355; AAA52454.1; -
CC PIR; A32947; A32947.
CC Genbank; HGNC:3748; FLG.
CC MIM; 135940;
CC GO; GO:0005882; C:intermediate filament; NAS.
CC GO; GO:0005198; F:structural molecule activity; NAS.
CC GO; GO:0007275; P:development; NAS.
CC InterPro; IPR003303; Filaggrin.

DR Pfam; PF03516; Filaggrin; 2.
 DR PRINTS; PR00487; FILAGRIN.
 KW Phosphorylation; Cytullination; Developmental protein.
 FT NON TER 1
 SQ SEQUENCE 416 AA; 44105 MW; DEEA3218BA043F32 CRC64;

Query Match 68.7%; Score 79; DB 1; Length 416;
 Best Local Similarity 80.0%; Pred. No. 0.00012;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHOESTXGRSGRSGSGS 21
 Db 6 QSHOESTXGRSGRSGSGS 25

RESULT 2

VE2_HPV47 STANDARD; PRT; 506 AA.
 AC P22420;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.
 GN E2.
 OS Human papillomavirus type 47.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 CX NCBI_TaxID=10594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9028161; PubMed=2162112;
 RA Kiyono T., Adachi A., Ishibashi M.;
 RT "Genome organization and taxonomic position of human papillomavirus
 type 47 inferred from its DNA sequence.";
 RL Virology 177:401-405(1990).

CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2E RESPONSE ELEMENT (5'-ACGNNNNNGT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2E'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY SERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M33305; AAA46979.1; -
 DR PIR; D35324; W2ML47.
 DR HSSP; P03122; 280P.
 DR InterPro; IPR000427; E2_C.
 DR InterPro; IPR001866; E2_N.
 DR InterPro; IPR009021; Viral_DNA_bd.
 DR Pfam; PF00511; E2_C; 1.
 DR Pfam; PF00508; E2_N; 1.
 DR ProDom; PD000672; E2_C; 1.
 DR ProDom; PD000678; E2_N; 1.
 DR Early protein; Transcription regulation; Activator; DNA-binding;
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 506 AA; 57478 MW; 92C37F4BF825065 CR654;

Query Match 44.3%; Score 51; DB 1; Length 506;
 Best Local Similarity 66.7%; Pred. No. 2.1;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 7 STXGRSGRSGRSGS 21

Db 346 NTRGRGRGRGRGSGS 360

RESULT 3

AD11_MOUSE STANDARD; PRT; 773 AA.

AC Q9RIV4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ADAM 11 precursor (A disintegrin and metalloproteinase domain 11)
 DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)
 GN (MDC).
 GN ADAM11 OR MDC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99365303; PubMed=10433968;
 RA Sagane K., Yamazaki K., Mizui Y., Tanaka I.;
 RT "Cloning and chromosomal mapping of mouse ADAM11, ADAM22 and ADAM23.";
 RL Gene 236:79-86(1999).

CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non
 CC catalytic metalloproteinase-like protein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in the brain. Weakly
 CC detected in the heart, liver and testis.
 CC -1- DOMAIN: A conserved motif [AVN(E/D)CD] within the disintegrin-like
 CC domain could be involved in the binding to the integrin receptor.
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 disintegrin domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB009676; BAA83384.1; -
 DR HSSP; P18619; 1EVL.
 DR MEROPS; M12.976; -
 DR MGD; MGI:1096667; Adam11.
 DR InterPro; IPR006586; ADAM cysteine.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006025; Pept_M_zn_BS.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR Pfam; PF00200; disintegrin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Repolysein; 1.
 DR PRINTS; PR00289; DISINTEGRIN.
 DR ProDom; PD000664; Disintegrin; 1.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00050; DISIN; 1.
 DR PROSITE; PS50215; ADAM_MEROP; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS50214; DISINTEGRIN_2; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_3; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain.
 FT SIGNAL 1
 POTENTIAL. 24

```

FT PROPEP 25 229 BY SIMILARITY.
FT CHAIN 230 773 ADAM 11.
FT DOMAIN 230 738 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 739 759 POTENTIAL.
FT DOMAIN 760 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 230 442 METALLOPROTEASE-LIKE.
FT DOMAIN 448 535 DISINTEGRIN-LIKE.
FT DOMAIN 536 680 EGF-LIKE.
FT DOMAIN 681 713 BY SIMILARITY.
FT DISULFID 507 520 POTENTIAL.
FT DISULFID 681 696 BY SIMILARITY.
FT DISULFID 690 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 773 AA; 84134 MW; 9A8A125FB3F743DD CRC64;

Query Match 43.0%; Score 49.5; DB 1; Length 773;
Best Local Similarity 45.0%; Pred. No. 5.4;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

3 CHO2-STXGRSGRCGRSGS 21
571 CTEKLNWETGRNGCRKGS 590

RESULT 4
AD1_HUMAN STANDARD; PRT; 769 AA.
ID AD1_HUMAN STANDARD; PRT; 769 AA.
AC 075078; Q14808; Q14809; Q14810.
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAM 11 precursor (A disintegrin-like and metalloproteinase domain 11)
DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)
DE (MDC)
GN ADAM11 OR MDC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=98359734; PubMed=693107;
RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.,
RA "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2
RT and MDC3: novel human cellular disintegrins highly expressed in the
RT brain.";
RL Biochem. J. 334:93-98(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Cerebellum;
RX MEDLINE=94073190; PubMed=8252040;
RA Emi M., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I.,
RA Kasumi F., Nakamura Y.;
RT "A novel metalloproteinase/disintegrin-like gene at 17q21.3 is
RT somatically rearranged in two primary breast cancers.";
RL Nat. Genet. 5:151-157(1993).
RN [3]
RP SEQUENCE OF 106-769 FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=Brain, Breast, Ovary, and Testis;
RX MEDLINE=95044425; PubMed=7956356;
RA Katagiri T., Harada Y., Emi M., Nakamura Y.;
RT "Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron
RT organization and alternative splicing.";
RL Cytogenet. Cell Genet. 68:39-44(1995).
CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non
CC catalytic metalloproteinase-like protein.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

```

```

CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long; Synonyms=MDC-769;
CC Name=Short; Synonyms=MDC-524;
CC IsoId=O75078-1; Sequence=Displayed;
CC IsoId=O75078-2; Sequence=VSP_005472, VSP_005473, VSP_005474,
CC VSP_005475;
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain. Slightly
CC detected or not at all in other tissues.
CC -1- DOMAIN: A conserved motif [AVNIE/D/CD] within the disintegrin-like
CC domain could be involved in the binding to the integrin receptor.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- DISEASE: Rearrangements occur in breast and ovarian cancers which
CC involve multiple exons and disrupt the coding region.
CC -1- SIMILARITY: Belongs to peptidase family M12b.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 disintegrin domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009675; BAA32352.1; -.
DR EMBL; D17390; BAA04213.1; -.
DR EMBL; D31872; BAA06670.1; -.
DR EMBL; D31872; BAA06671.1; -.
DR PIR; I65967; I65967.
DR HSP; P18619; IFLV.
DR MEROPS; M12.976; -.
DR Genew; HGNC:189; ADAM11.
DR MIM; 155120; -.
DR GO; GO:0005886; C:Plasma membrane; TAS.
DR GO; GO:0005178; F:integrin binding; TAS.
DR GO; GO:0008237; F:metalloproteinase activity; TAS.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006025; Pept_M_2n_BS.
DR InterPro; IPR001590; Peptidase_M12b.
DR InterPro; IPR002870; Peptidase_M12b_N.
DR Pfam; PF00200; disintegrin; 1.
DR Pfam; PF01562; pep_M12b_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00289; DISINTEGRIN.
DR PRODOM; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS50215; ADAM_MEROP; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR Signal; Glycoprotein; Transmembrane; EGF-like domain;
DR Alternative splicing.
DR SIGNAL 1 23
DR PROPEP 24 225 BY SIMILARITY.
DR CHAIN 226 769 ADAM 11.
DR DOMAIN 226 734 EXTRACELLULAR (POTENTIAL).
DR TRANSEM 735 755 POTENTIAL.
DR DOMAIN 756 769 CYTOPLASMIC (POTENTIAL).
DR DOMAIN 226 438 METALLOPROTEASE-LIKE.
DR DOMAIN 444 531 DISINTEGRIN-LIKE.
DR DOMAIN 532 676 CYS-RICH.
DR DOMAIN 677 709 EGF-LIKE.
DR DISULFID 349 433 BY SIMILARITY.

```

```

FT DISULFID 503 516 POTENTIAL.
FT DISULFID 677 692 BY SIMILARITY.
FT DISULFID 686 698 BY SIMILARITY.
FT DISULFID 700 709 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASAPLIC 1 99 Missing (in isoform Short).
FT VASAPLIC 100 104 /FtId=VSP_005472.
FT VASAPLIC 595 623 /FtId=VSP_005473.
FT VASAPLIC 595 623 DVLGFLCNISGAPFLDVLVDISSVT -> POCGRWV
FT VASAPLIC 595 623 LPLCOHLMSSARGPGGRHQ (in isoform Short).
FT VASAPLIC 624 769 /FtId=VSP_005474.
FT VASAPLIC 624 769 Missing (in isoform Short).
FT VASAPLIC 106 106 /FtId=VSP_005475.
FT VASAPLIC 325 325 Q -> H (IN REF. 1).
FT VASAPLIC 325 325 D -> N (IN REF. 2).
SQ SEQUENCE 769 AA; 83409 MW; 59B9C467BDDF36E CRC64;

Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred No. 7.6;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHOESTXGRSGRGRSGS 21
Db 567 CYEXLNEGTGRSGCRKGS 586

RESULT 5
AKP1_RAT STANDARD; PRT; 854 AA.
AC O88884; O88884;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE A kinase anchor protein 1, mitochondrial precursor (Protein kinase A
DE anchoring protein 1) (PKA1) (A-kinase anchor protein 121 kDa) (AKAP
DE 121) (Dual specificity A-kinase anchoring protein 1) (D-AKAP-1)
DE (Spermatid A-kinase anchor protein 84) (S-AKAP84).
GN AKAP1 OR AKAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Thyroid;
RX MEDLINE=98389771; PubMed=9722570;
RA Feliciello A.; Rubin C.S.; Avvedimento E.V.; Gottesman M.E.;
RT "Expression of a kinase anchor protein 121 is regulated by hormones in
RT thyroid and testicular germ cells.";
RL J. Biol. Chem. 273:23361-23366(1998).
CC -1- FUNCTION: BINDS TO TYPE I AND II REGULATORY SUBUNITS OF PROTEIN
CC KINASE A AND ANCHORS THEM TO THE CYTOPLASMIC FACE OF THE
CC MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Mitochondrial outer membrane (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=AKAP121;
CC IsoId=O88884-1; Sequence=Displayed;
CC Name=2; Synonyms=S-AKAP84;
CC IsoId=O88884-2; Sequence=VSP_002854, VSP_002855;
CC -1- TISSUE SPECIFICITY: testis specific.
CC -1- INDUCTION: BY THYROID STIMULATING HORMONE (TSH) AND CAMP OR CAMP-
CC ANALOG.
CC -1- DOMAIN: RII-alpha binding site, predicted to form an amphipathic
CC helix, could participate in protein-protein interactions with a
CC complementary surface on the R-subunit dimer.
CC -1- SIMILARITY: Contains 1 KH domain.
CC -1- SIMILARITY: Contains 1 Tudor domain.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL; AF068202; AAC61775.1; -
CC DR EMBL; AF092523; AAC61775.1; -
CC DR InterPro; IPR004087; KH dom.
CC DR InterPro; IPR004088; KH type 1.
CC DR InterPro; IPR004191; Maternal_tudor.
CC DR InterPro; IPR002999; Tudor.
CC DR Pfam; PF00013; KH; 1.
CC DR Pfam; PF00567; TUDOR; 1.
CC DR SMART; SM00322; KH; 1.
CC DR SMART; SM00333; TUDOR; 1.
CC DR PROSITE; PS00084; KH TYPE 1; 1.
CC DR PROSITE; PS00304; TUDOR; 1.
CC DR RNA-binding; Mitochondrion; Alternative splicing; Outer membrane;
CC Transmembrane; Transic peptide.
CC KW TRANSMEMBRANE 1 29 MITOCHONDRION (BY SIMILARITY).
CC FT CHAIN 30 854 A KINASE ANCHOR PROTEIN 1.
CC FT DOMAIN 303 316 PKA-RII SUBUNIT BINDING DOMAIN.
CC FT DOMAIN 558 622 KH.
CC FT DOMAIN 709 768 TUDOR.
CC FT VASAPLIC 523 544 GSDNSNDSDVSCGLTKPSP -> VAAPQGRHGRNGG
CC FT CAGFFEC (in isoform 2).
CC FT VASAPLIC 545 854 /FtId=VSP_002854.
CC FT VASAPLIC 545 854 Missing (in isoform 2).
CC FT CONFLICT 311 311 I -> T (IN REF. 1; AAC61775).
CC FT CONFLICT 439 439 V -> A (IN REF. 1; AAC61775).
CC SQ SEQUENCE 854 AA; 91447 MW; 6EFA30F8801A06E CRC64;

Query Match 41.3%; Score 47.5; DB 1; Length 854;
Best Local Similarity 47.8%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

QY 3 CHOESTXGRSGR-----CGRSG 20
Db 160 CKQESALGRTPGRGWSQCAASG 182

RESULT 6
FER2_AZOVI STANDARD; PRT; 106 AA.
AC P82802.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ferredoxin, 2Fe-2s (Shetna protein I) (Iron-sulfur protein I)
DE (2FeAVFdI).
OS Azotobacter vinelandii.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN (1)
RP SEQUENCE, SUBUNIT, AND MASS SPECTROMETRY.
RC STRAIN=DJ;
RX MEDLINE=99367931; PubMed=10439076;
RA Chatelet C.; Meyer J.J. (Shetna protein I) from Azotobacter
RT "The [2Fe-2S] protein I (Shetna protein I) from Azotobacter
RT vinelandii is homologous to the [2Fe-2S] ferredoxin from Clostridium
RT pasteurianum.";
RL J. Biol. Inorg. Chem. 4:311-317(1999).
CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS. MIGHT BE
CC INVOLVED IN NITROGEN FIXATION.
CC -1- COFACTOR: Binds 1 2Fe-2S cluster.
CC -1- SUBUNIT: Homodimer.

```

```
CC -1- MASS SPECTROMETRY: MW=22876; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE 2FE2S SHETNA-TYPE FERREDOXIN FAMILY.
DR HSSP; 066511; 1F37.
KW Electron transport; Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KM Nitrogen fixation.
FT METAL 10 10 IRON-SULFUR (2FE-2S).
FT METAL 23 23 IRON-SULFUR (2FE-2S).
FT METAL 55 55 IRON-SULFUR (2FE-2S).
FT METAL 59 59 IRON-SULFUR (2FE-2S).
SQ SEQUENCE 106 AA; 11264 MW; 803BB1158A96CB67 CRC64;

Query Match 40.0%; Score 46; DB 1; Length 106;
Best Local Similarity 42.1%; Pred. No. 2.6;
Matches 8; Conservative 1; Mismatches 10; Gaps 0;

QY 3 CHOESTXGRSGRCGRSGS 21
DB 10 CAQNPBAGHPRGSCGAKGA 28

RESULT 7
ID 2205 HUMAN STANDARD; PRT; 504 AA.
AC 095201;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 205 (Zinc finger protein 210).
GN ZNF205 OR ZNF210.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=99005537; Pubmed=9787081;
RA Deng Z., Centola M., Chen X., Sood R., Vedula A., Fischel-Ghodsian N.,
RA Kastner D.L.;
RT "Identification of two Kruppel-related zinc finger genes (ZNF200 and
RT ZNF210) from human chromosome 16p13.3."
RL Genomics 53:97-103(1998).
CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION REPRESSOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, SKELETAL MUSCLE, PANCREAS
CC AND BRAIN. WEAKLY EXPRESSED IN PLACENTA, LUNG, LIVER, KIDNEY AND
CC THYRUS.
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF060865; AAC70007.1; -.
DR HSSP; P08046; 1A1G.
DR Genew; HGNC:12966; ZNF205.
DR MIM; 603436; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003677; F:DNA binding; NAS.
DR GO; GO:0008270; F:zinc ion binding; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf_C2H2; 8.
DR Prodom; PD000003; Znf_C2H2; 8.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 8.
```

```
DR PROSITE; PS50905; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KM Nuclear protein; Repressor; Repeat.
FT DOMAIN 74 143 KRAB.
FT ZN_FING 258 280 C2H2-TYPE.
FT ZN_FING 286 308 C2H2-TYPE.
FT ZN_FING 314 336 C2H2-TYPE.
FT ZN_FING 342 364 C2H2-TYPE.
FT ZN_FING 370 392 C2H2-TYPE.
FT ZN_FING 398 420 C2H2-TYPE.
FT ZN_FING 426 448 C2H2-TYPE.
FT ZN_FING 454 476 C2H2-TYPE.
SQ SEQUENCE 504 AA; 55312 MW; 33C0430D748ABF8C CRC64;

Query Match 39.1%; Score 45; DB 1; Length 504;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 QCHOESTXGRSGRCGRSG 20
DB 184 QCAQEPACGRSGSPAKDSG 202

RESULT 8
ID 5E5 RAT STANDARD; PRT; 825 AA.
AC Q65003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 5E5 antigen.
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=wislar; TISSUE=Brain;
RX MEDLINE=9601519; Pubmed=8537300;
RA Suzuki E., Kojima N., Yoshimura K., Iyemura K., Obara K., Akagawa K.;
RT "Cloning and sequence analysis of cDNA for a possible DNA-binding
RT protein 5E5 in the nervous system."
RL J. Biochem. 118:122-128(1995).
CC -1- FUNCTION: MIGHT HAVE DNA-BINDING ABILITY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Expressed in neurons.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D37934; BAA07153.1; -.
DR PIR; JC4163; JC4163.
KW DNA-binding; Nuclear protein; Antigen.
SQ SEQUENCE 825 AA; 86831 MW; AF67FE2PD555BDF CRC64;

Query Match 38.3%; Score 44; DB 1; Length 825;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 GRSRGRCGRSGS 21
DB 589 GRGRGRCGRSGS 600

RESULT 9
M2GD_HUMAN
```

ID M2GD HUMAN STANDARD; PRT; 866 AA.
 AC O9U117;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 15-methylglycine dehydrogenase, mitochondrial precursor (EC 1.5.99.2)
 DE (ME2G.DYD).
 GN DMGDH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A., AND VARIANTS PRO-279; ALA-530 AND SER-646.
 RX MEDLINE=20232164; PubMed=10767172;
 RA Binzak B.A., Vockley J.G., Jenkins R.B., Vockley J.;
 RT "Structure and analysis of the human dimethylglycine dehydrogenase
 gene";
 RL Mol. Genet. Metab. 69:181-187(2000).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT DMGDH ARG-109.
 RX MEDLINE=21152272; PubMed=11231903;
 RA Binzak B.A., Meyers R.A., Moolenaar S.H., Lee Y.-M., Hwu W.-L.,
 RA Poggi-Bach J., Engelke U.F.H., Hoard H.M., Vockley J.G., Vockley J.;
 RT "Cloning of dimethylglycine dehydrogenase and a new human inborn error
 of metabolism, dimethylglycine dehydrogenase deficiency";
 RL Am. J. Hum. Genet. 68:839-847(2001).
 RN [3]
 RP DISEASE.
 RX MEDLINE=99203382; PubMed=10102904;
 RA Moolenaar S.H., Poggi-Bach J., Engelke U.F.H., Corstiaensen J.M.B.,
 RA Heerschap A., de Jong J.G.N., Binzak B.A., Vockley J., Meyers R.A.;
 RT "Defect in dimethylglycine dehydrogenase, a new inborn error of
 metabolism: NMR spectroscopy study";
 RL Clin. Chem. 45:459-464(1999).
 CC -1- CATALYTIC ACTIVITY: N,N-dimethylglycine + acceptor + H(2)O =
 sarcosine + formaldehyde + reduced acceptor.
 CC -1- COFACTOR: FAD; covalently linked.
 CC -1- PATHWAY: Choline metabolism.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- DISEASE: Defects in DMGDH are the cause of DMGDH deficiency
 (DMGDH) [MIM:605850]. DMGDH is a disorder characterized by fish
 odor, muscle fatigue with increased serum creatine kinase.
 CC Biochemically it is characterized by an increase of N,N-
 dimethylglycine (DMG) in serum and urine.
 CC -1- SIMILARITY: Belongs to the gcvT family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
 or send an email to license@isb-sb.ch).
 CC EMBL; AF11858; AAF21941.1; -
 DR MIM; 605849; -
 DR MIM; 605850; -
 DR GO; GO:0005739; C:mitochondrion; TAS.
 DR GO; GO:0008246; F:electron transfer flavoprotein; TAS.
 DR InterPro; IPR006076; FAD_oxred.
 DR InterPro; IPR006222; GCV_T.
 DR InterPro; IPR000205; NAD_BS.
 DR Pfam; PF01266; DAO; 1.
 DR Pfam; PF01571; GCV_T; 1.
 KM Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide;
 KM Disease mutation; Polymorphism.
 KM TRANSIT 1 50
 FT CHAIN 51 866
 FT BINDING 91 91
 FT VARIANT 109 109
 H -> R (in DMGDH).
 /FTID=VAR_011505.

FT VARIANT 279 279 S -> P (in dbSNP:532964).
 FT FTID=VAR_014950.
 FT VARIANT 530 530 G -> A (in dbSNP:1805073).
 FT FTID=VAR_014951.
 FT VARIANT 646 646 P -> S (in dbSNP:1805074).
 FT FTID=VAR_014952.
 SQ SEQUENCE 866 AA; 96806 MW; 546E1F6C1656AC04 C64;
 Query Match 38.3%; Score 44; DB 1; Length 866;
 Best Local Similarity 44.4%; Pred. No. 40;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 3 CHESTXGRSGRCGRSG 20
 DB 17 CPGCSPRPRSVCGREG 34
 RESULT 10
 ID DEF1 MYTGA STANDARD; PRT; 38 AA.
 AC P80571;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Defensin MGD-1.
 GN FH3.
 OS Mytilus galloprovincialis (Mediterranean mussel).
 CC Eukaryota; Metazoa; Mollusca; Bivalvia; Pelecimorpha; Mytiloidea;
 CC Mytiloidea; Mytilidae; Mytilus.
 CC NCBI_TaxID=29158;
 RN [1]
 RP SEQUENCE.
 RX TISSUE=Hemolymph;
 RX MEDLINE=97025339; PubMed=8925841;
 RA Hubert F., Noel T., Roch P.;
 RT "A member of the arthropod defensin family from edible Mediterranean
 mussels (Mytilus galloprovincialis).";
 RL Eur. J. Biochem. 240:302-306(1996).
 RN [2]
 RP ERRATUM.
 RA Hubert F., Noel T., Roch P.;
 RL Eur. J. Biochem. 240:815-815(1996).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20341401; PubMed=11087396;
 RA Yang Y.S., Mita G., Chavanien A., Calas B., Sanchez J.F., Roch P.,
 RA Amelias A.;
 RT "Solution structure and activity of the synthetic four-disulfide bond
 Mediterranean mussel defensin (MGD-1).";
 RL Biochemistry 39:14436-14447(2000).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the arthropod defensin family. Subfamily 2.
 CC PIR; S74088; S74088.
 DR PDB; 1FJN; 20-DEC-00.
 DR InterPro; IPR001542; Defensin_annot.
 DR Pfam; PF01097; Arthro_defensin; 1.
 DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
 KM Antibiotic; 3D-structure.
 FT DISULFID 4 25
 FT DISULFID 10 33
 FT DISULFID 14 35
 FT DISULFID 21 38
 SQ SEQUENCE 38 AA; 4275 MW; 6F1C0C0CF0E69E76 C64;
 Query Match 37.4%; Score 43; DB 1; Length 38;
 Best Local Similarity 47.4%; Pred. No. 2; 7;
 Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;
 QY 1 HQCHQ--ESTYGRSGRCG 17
 DB 8 YQCHRRKSLPRGCGYCG 26


```

RESULT 11
CAL4 MOUSE STANDARD; PRT; 1669 AA.
ID _CAL4 MOUSE
AC P02463;
DT 21-JUL-1986 (Rel. 02, Created)
DT 01-FEB-1992 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Collagen alpha 1(IV) chain precursor.
GN COL4A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89197932; PubMed=2703490;
RA Muthukumaran G., Blumberg B., Kurkinen M.;
RT "The complete primary structure for the alpha 1-chain of mouse
RT collagen IV. Differential evolution of collagen IV domains.";
RL J. Biol. Chem. 264:6310-6317(1989).
RN [2]
RP SEQUENCE OF 1-1154 FROM N.A.
RX MEDLINE=88112221; PubMed=3338568;
RA Wood L., Theriault N., Vogeli G.;
RT "cDNA clones completing the nucleotide and derived amino acid
RT sequence of the alpha 1 chain of basement membrane (type IV) collagen
RT from mouse.";
RL FEBS Lett. 227:5-8(1988).
RN [3]
RP SEQUENCE OF 1149-1424 FROM N.A.
RX MEDLINE=86301886; PubMed=3755692;
RA Nath P., Laurent M., Horn E., Sobel M.E., Zon G., Vogeli G.;
RT "Isolation of an alpha 1 type-IV collagen cDNA clone using a
RT synthetic oligodeoxynucleotide.";
RL Gene 43:301-304(1986).
RN [4]
RP SEQUENCE OF 1276-1669 FROM N.A.
RX MEDLINE=85127033; PubMed=2578961;
RA Oberhaeumer I., Laurent M., Schwarz U., Sakurai Y., Yamada Y.,
RA Vogeli G., Voss T., Siebold B., Glanville R.W., Kuhn K.;
RT "Amino acid sequence of the non-collagenous globular domain (NC1) of
RT the alpha 1(IV) chain of basement membrane collagen as derived from
RT complementary DNA.";
RL Eur. J. Biochem. 147:217-224(1985).
RN [5]
RP SEQUENCE OF 1441-1669 FROM N.A.
RX MEDLINE=87250460; PubMed=3597383;
RA Kurkinen M., Condon W.R., Blumberg B., Barlow D., Quinones S.,
RA Saus J., Pihlajaniemi T.;
RT "Extensive homology between the carboxyl-terminal peptides of mouse
RT alpha 1(IV) and alpha 2(IV) collagen.";
RL J. Biol. Chem. 263:8496-8499(1987).
RN [6]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=86196099; PubMed=3009468;
RA Sakurai Y., Sullivan M., Yamada Y.;
RT "Alpha 1 type IV collagen gene evolved differently from fibrillar
RT collagen genes.";
RL J. Biol. Chem. 261:6654-6657(1986).
RN [7]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89066738; PubMed=3198626;
RA Kayes P., Wood L., Theriault N., Kurkinen M., Vogeli G.;
RT "Head-to-head arrangement of murine type IV collagen genes.";
RL J. Biol. Chem. 263:19274-19277(1988).
RN [8]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89071759; PubMed=3200851;
RA Burdalo P.D., Martin G.R., Yamada Y.;
RT "Alpha 1(IV) and alpha 2(IV) collagen genes are regulated by a
RT bidirectional promoter and a shared enhancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:9679-9682(1988).
RN [9]

```

```

RP SEQUENCE OF 1-129 FROM N.A.
RX MEDLINE=88243724; PubMed=379041;
RA Killen P.D., Burdalo P., Sakurai Y., Yamada Y.;
RT "Structure of the amino-terminal portion of the murine alpha 1(IV)
RT collagen chain and the corresponding region of the gene.";
RL J. Biol. Chem. 263:8706-8709(1988).
CC -1- FUNCTION: Type IV collagen is the major structural component of
CC glomerular basement membranes (GBM), forming a 'chicken-wire'
CC network together with laminins, proteoglycans and enactin/
CC nidogen.
CC -1- SUBUNIT: There are six type IV collagen isoforms, alpha 1(IV)-
CC alpha 6(IV), each of which can form a triple helix structure with
CC 2 other chains to generate type IV collagen network.
CC -1- DOMAIN: Alpha chains of type IV collagen have a noncollagenous
CC domain (NC1) at their C-terminus, frequent interruptions of the G-
CC X-Y repeats in the long central triple-helical domain (which may
CC cause flexibility in the triple helix), and a short N-terminal
CC triple-helical 7S domain.
CC -1- PTM: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -1- PTM: Type IV collagens contain numerous cysteine residues which
CC are involved in inter- and intramolecular disulfide bonding. 12 of
CC these, located in the NC1 domain, are conserved in all known type
CC IV collagens.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annotation/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03758; AAA37439.1; -
DR EMBL; M23333; AAA51625.1; -
DR EMBL; J04694; AAA50292.1; -
DR EMBL; X06777; CAA28946.1; -
DR EMBL; X02201; CAA26132.1; -
DR EMBL; M15832; AAA37340.1; -
DR EMBL; M14042; AAA37342.1; -
DR EMBL; M12879; AAA37343.1; -
DR EMBL; M13024; -; NOT_ANNOTATED_CDS.
DR EMBL; M13025; -; NOT_ANNOTATED_CDS.
DR EMBL; M13026; AAA37344.1; -
DR EMBL; M13027; AAA37345.1; -
DR EMBL; M13043; AAA37346.1; -
DR EMBL; J04448; AAA37437.1; -
DR PIR; A33585; CGMS4B.
DR MGB; MGI:88454; Col4a1.
DR GO; GO:0005604; C:Basement membrane; IDA.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001442; Procollagen4_C.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 23.
DR ProDom; PD000007; C1g_helix; 6.
DR ProDom; PD003923; ProcollagenC4; 1.
DR SMART; SM00111; C4; 2.
KW Extracellular matrix; Connective tissue; Basement membrane;
KW Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
FT SIGNAL 1 27
FT PROPEP 28 172
FT CHAIN 173 1669
FT DOMAIN 173 1440
FT TRIPLE-HELICAL REGION.
FT 1441 1669 NONHELICAL REGION (NC1).
FT DISULFD 1460 1551 OR 1548 (BY SIMILARITY).
FT DISULFD 1493 1548 OR 1551 (BY SIMILARITY).
FT DISULFD 1505 1511 BY SIMILARITY.
FT DISULFD 1570 1665 OR 1662 (BY SIMILARITY).
FT DISULFD 1604 1662 OR 1665 (BY SIMILARITY).
FT DISULFD 1616 1622 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 26 26 A -> P (IN REF. 2).

```

FT CONFLICT 186 186 S -> L (IN REF. 2).
 FT CONFLICT 319 319 Q -> S (IN REF. 2).
 FT CONFLICT 369 369 L -> L (IN REF. 2).
 FT CONFLICT 403 403 L -> F (IN REF. 2).
 FT CONFLICT 481 481 P -> L (IN REF. 2).
 FT CONFLICT 493 493 Q -> H (IN REF. 2).
 FT CONFLICT 712 712 S -> I (IN REF. 2).
 FT CONFLICT 813 813 E -> Q (IN REF. 2).
 FT CONFLICT 982 982 Q -> H (IN REF. 2).
 FT CONFLICT 1397 1397 V -> S (IN REF. 3).
 SQ SEQUENCE 1669 AA; 160680 MW; 42916591E52058E9 CRC64;

Query Match 37.4%; Score 43; DB 1; Length 1669;
 Best Local Similarity 41.2%; Pred. No. 1.1e+02;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Dy 4 H0ESTXGRRCRCRSG 20
 19 HERSRAMAKDCGSG 35

RESULT 12

AD21_MOUSE STANDARD; PRT; 729 AA.

AC Q90T76;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAM 21 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase domain 21) (ADAM 31).
 GN ADAM21 OR ADAM31.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Bone marrow;
 RX MEDLINE=20289099; PubMed=10830287;
 RA Liu L., Smith J.W.;
 RT "Identification of ADAM 31: a protein expressed in Leydig cells and specialized epithelia."
 RL Endocrinology.141:2033-2042(2000).
 [2]

RE SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleishmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Bash G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.U., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima U., Mazzarelli J., Mombauts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: May be involved in sperm maturation and/or
 fertilization. May also be involved in epithelia functions
 associated with establishing and maintaining gradients of ions or
 nutrients.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (potential).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Highly expressed in Leydig cells. Expressed
 also in cauda epididymidis, vas deferens, convoluted tubules,
 kidney and the parietal cells of stomach. Not detected on
 developing spermatocytes or mature sperm.
 CC -1- DOMAIN: A tripeptide motif (VGE) within disintegrin-like domain
 could be involved in the binding to egg integrin receptor and thus
 could mediate sperm/egg binding.
 CC -1- DOMAIN: The cysteine-rich domain encodes putative cell-fusion
 peptides, which could be involved in sperm-egg fusion.
 CC -1- PTM: Has no obvious cleavage site for furin endopeptidase,
 suggesting that the proteolytic processing is regulated (By
 similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 disintegrin domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF251559; AAF74731.1; -;
 CC EMBL; AK014827; BAB29569.1; -;
 CC HSSP; P18619; 1FVL.
 CC MEROPS; M12.233; -;
 CC MGD; MGI:1861229; Adam21.
 CC InterPro; IPR005586; ADAM cysteine.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR001818; Pept_M10A_M12B.
 CC InterPro; IPR006025; Pept_M2n_S8.
 CC InterPro; IPR001590; peptidase_M12B.
 CC InterPro; IPR002870; peptidase_M12B_N.
 CC Pfam; PF00200; disintegrin; 1.
 CC Pfam; PF01562; pep_M12B_propep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC PRINTS; PRO0289; DISINTEGRIN.
 CC PRODOM; PD000664; Disintegrin; 1.
 CC SMART; SM00608; ACR; 1.
 CC SMART; SM0050; DISIN; 1.
 CC SMART; SM0181; EGF; 1.
 CC PROSITE; PS50215; ADAM_MEROP; 1.
 CC PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 CC PROSITE; PS00427; DISINTEGRIN 1; FALSE NEG.
 CC PROSITE; PS50214; DISINTEGRIN 2; 1.
 CC PROSITE; PS00022; EGF_1; FALSE_NEG.
 CC PROSITE; PS01185; EGF_2; 1.
 CC PROSITE; PS50026; EGF_3; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC HYDROLASE; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Transmembrane; EGF-like domain.
 FT SIGNAL 1 39
 FT PROPEP 40 209
 FT CHAIN 210 729
 FT DOMAIN 210 685
 FT TRANSMEM 686 706
 FT DOMAIN 707 729
 FT DOMAIN 210 402
 FT DOMAIN 410 496
 FT DOMAIN 497 639
 FT DOMAIN 638 667
 FT SITE 178 178
 FT SITE 345 345
 FT ACT SITE 346 346
 FT METAL 349 349
 FT METAL 355 355
 FT DISULFID 320 397
 FT DISULFID 360 382
 FT DISULFID 468 481
 FT POTENTIAL.

```

FT DISULFID 638 649 BY SIMILARITY.
FT DISULFID 643 655 BY SIMILARITY.
FT DISULFID 657 666 BY SIMILARITY.
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 729 AA; 80849 MW; E70B79BC46B4549 CRC64;

Query Match 37.0%; Score 42.5; DB 1; Length 729;
Best Local Similarity 36.4%; Pred. No. 56;
Matches 8; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

OY 1 HQCHOE-STXGRSGRGRSGS 21
DB 531 HNCVKEINLQGNRFCHCCTDGT 552

RESULT 13
YICG_CABEL STANDARD; PRT; 102 AA.
AC 08MNT7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Hypothetical UPF0287 protein C35D10.17 in chromosome III.
GN C35D10.17.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN REVISONS.
RA Waterston R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the UPF0287 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U21324; AAM2045.1. -
CC WormPep; C35D10.17; C350625.
CC Hypothetical protein.
CC SEQUENCE 102 AA; 11993 MW; C519494F4098A23D CRC64;

Query Match 36.5%; Score 42; DB 1; Length 102;
Best Local Similarity 40.0%; Pred. No. 9; 8;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 QCHQESTXGRSGRGRG 16
DB 23 RCHSEKPIGKMGKGC 37

RESULT 14
LSM4_CABEL STANDARD; PRT; 123 AA.
ID LSM4_CABEL
AC Q19952;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable U6 snRNA-associated Sm-like protein LSM4.

```

```

GN F32A5.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Paulley A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Binds specifically to the 3'-terminal U-tract of U6
CC snRNA (By similarity).
CC -1- SUBUNIT: Lsm subunits form a heteromer with a doughnut shape (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U20864; AAC4661.1. -
CC PIR; T16234; T16234.
CC WormPep; F32A5.7; CE01277.
CC InterPro; IPR006649; snRNP.
CC InterPro; IPR001163; snRNP_Sm.
CC Pfam; PF01423; LSM; 1.
CC ProDom; PD020287; snRNP; 1.
CC SMART; SM00651; Sm; 1.
CC Nuclear protein, Ribonucleoprotein; mRNA splicing; mRNA processing;
CC RNA-binding.
CC SEQUENCE 123 AA; 13593 MW; B3EB874B67705109 CRC64;

Query Match 36.5%; Score 42; DB 1; Length 123;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 2 QCHQESTXGRSGRGRG 20
DB 90 QCHQESTXGRSGRGRG 108

RESULT 15
Y4CG_RHISN STANDARD; PRT; 305 AA.
ID Y4CG_RHISN
AC P55389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE Probable DNA-invertase Y4CG.
GN Y4CG.
OS Rhizobium sp. (strain NGR34).
OG Plasmid sym DNGR234.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_Taxid=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; Pubmed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: Belongs to the site-specific recombinase resolvase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000068; AAB92422.1; -.
DR HSSP; P03012; 2RSL.
DR InterPro; IPR006118; Recombinase.
DR InterPro; IPR006119; resolvase_N.
DR Pfam; PF00239; resolvase; 1.
DR PROSITE; PS00397; RECOMBINASES_1; 1.
DR PROSITE; PS00398; RECOMBINASES_2; 1.
DR Hypothetical protein; DNA recombination; DNA integration; DNA-binding;
KW DNA invertase; Plasmid.
FT ACT_SITE 23
FT TRANSIENT COVALENT LINKAGE TO DNA DURING
FT STRAND CLEAVAGE AND REJOINING
FT (BY SIMILARITY).
SQ SEQUENCE 305 AA; 34277 MW; 48C03BD3AA9420F CRC64;
FT
OY 1 HOCH0ESTXGRGR 15
Db 42 HRIHQHSGASRAR 56

```

```

Query Match 36.5%; Score 42; DB 1; Length 305;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

Search completed: September 28, 2004, 06:05:53
 Job time : 8.98438 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 44.4062 Seconds

(without alignments)
149.211 Million cell updates/sec

Title: US-09-308-150-10

Perfect score: 115
Sequence: 1 HQCHQESTXGRSGRCGRSGS 21

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mbc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	78.3	797	4	Q16824	Q16824 homo sapien
2	86	74.8	990	4	Q15206	Q15206 homo sapien
3	86	74.8	1218	4	Q05331	Q05331 homo sapien
4	79	68.7	798	4	Q05403	Q05403 homo sapien
5	79	68.7	1084	4	Q01212	Q01212 homo sapien
6	70	60.9	591	4	Q01720	Q01720 homo sapien
7	70	60.9	687	4	Q05402	Q05402 homo sapien
8	65	56.5	465	4	Q03838	Q03838 homo sapien
9	64	55.7	322	4	Q75370	Q75370 homo sapien
10	53.5	46.5	305	11	Q08C86	Q08C86 mus musculu
11	53	46.1	113	11	Q0CZ97	Q0CZ97 mus musculu
12	53	46.1	113	11	Q0BT22	Q0BT22 mus musculu
13	53	46.1	130	11	Q0CDR9	Q0CDR9 mus musculu
14	53	46.1	822	3	Q0P312	Q0P312 neurospora
15	53	46.1	947	12	Q0B489	Q0B489 dengue viru
16	52	45.2	1831	5	Q9W2E1	Q9W2E1 drosophila

17	52	45.2	1872	5	Q7YU83	Q7YU83 drosophila
18	50	43.5	214	16	Q8FLQ3	Q8FLQ3 corynebacte
19	49.5	43.0	305	11	Q0DCV4	Q0DCV4 mus musculu
20	49.5	43.0	778	11	Q7YU87	Q7YU87 mus musculu
21	49	42.6	678	10	Q0R200	Q0R200 oryza sativ
22	48	41.7	214	10	Q08N84	Q08N84 arabidopsis
23	48	41.7	254	10	Q08MP2	Q08MP2 arabidopsis
24	47.5	41.3	305	11	Q0C0Y9	Q0C0Y9 mus musculu
25	47.5	41.3	302	3	Q073D9	Q073D9 neurospora
26	47	40.9	157	10	Q9HBN1	Q9HBN1 homo sapien
27	47	40.9	128	10	Q0ASD7	Q0ASD7 oryza sativ
28	47	40.9	266	10	Q06410	Q06410 daucus caro
29	47	40.9	266	10	Q06411	Q06411 daucus caro
30	47	40.9	451	15	Q08B08	Q08B08 avian leuko
31	47	40.9	491	15	Q08B05	Q08B05 avian leuko
32	47	40.9	491	15	Q08B07	Q08B07 avian leuko
33	47	40.9	604	15	Q08132	Q08132 avian myelo
34	47	40.9	730	2	Q02301	Q02301 pseudomonas
35	47	40.9	1461	11	Q09TLP3	Q09TLP3 mus musculu
36	46.5	40.4	77	10	Q083P2	Q083P2 oryza sativ
37	46.5	40.4	338	16	Q09R1L1	Q09R1L1 deinococcus
38	46.5	40.4	443	10	Q09AR06	Q09AR06 oryza sativ
39	46	40.0	68	8	Q05WS9	Q05WS9 pinus korai
40	46	40.0	101	16	Q7VQD5	Q7VQD5 candidatus
41	46	40.0	170	10	Q04LT6	Q04LT6 oryza sativ
42	46	40.0	170	10	Q07CH7	Q07CH7 oryza sativ
43	46	40.0	268	10	Q06408	Q06408 daucus caro
44	46	40.0	644	16	Q08E32	Q08E32 bradyrhizob
45	46	40.0	849	10	Q0C7M2	Q0C7M2 arabidopsis

ALIGNMENTS

RESULT 1
ID Q16824 PRELIMINARY; PRT; 797 AA.

AC Q16824; 01-NOV-1996 (TREMURel. 01, Created)
DT 01-NOV-1996 (TREMURel. 01, Last sequence update)
DT 01-JUN-2003 (TREMURel. 24, Last annotation update)
DE Profilaggrin (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91064347; PubMed=2248957;
RX GAN S.O., McBride O.W., Idler W.W., Markova N., Steiner P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814]."
RL Biochemistry 29:9432-9440(1990).
RL EMBL; M60502; AAA63248.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003303; Flaggrin.
DR Pfam; PF03516; Flaggrin; 4.
DR PRINTS; PRO0487; FLAGGRIN.
FT NON_TER
SQ SEQUENCE 797 AA; 85176 MW; 60E6184763BDA66B CRC64;

Query Match 78.3%; Score 90; DB 4; Length 797;
Best Local Similarity 81.0%; Pred. No. 8.8e-06;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 HQCHQESTXGRSGRCGRSGS 21
DB 425 HESHQESTXGRSGRCGRSGS 445

RESULT 2

```

015206 ID 015206 PRELIMINARY; PRT; 990 AA.
AC Q15206;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Profilaggrin (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=9125119; PubMed=2043621;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene.";
RL Biochemistry 30:5814-5814 (1991).
DR EMBL: M60494; AAA63244.1; -
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF03516; Filaggrin; 6.
DR PRINTS: PR00487; Filaggrin.
FT NON TER 990
SQ SEQUENCE 990 AA; 106453 MW; A8396F10F6A91991 CRC64;

Query Match 74.8%; Score 86; DB 4; Length 990;
Best Local Similarity 81.0%; Pred. No. 4,6e-05;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HOCHBESTXGRSGRCGRSGS 21
DB 225 HOCHBESTXGRSGRCGRSGS 245

RESULT 3
Q05331 ID 005331 PRELIMINARY; PRT; 1218 AA.
AC 005331;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE FILAGGRIN (PROFILAGGRIN) (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FORESKIN;
RX MEDLINE=93109348; PubMed=8417356;
RA Markova N.G., Marekov L.N., Chipev C.C., Gan S.Q., Idler W.W.,
RA Steinert P.M.;
RT "Profilaggrin is a major epidermal calcium-binding protein.";
RL Mol. Cell. Biol. 13:613-623 (1993).
RN [2]
RP FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
RN DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
RN TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
CC 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &

```

```

CC PROTEOLYTICALLY CLEAVED.
CC -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
CC REPEATS.
DR EMBL: M96943; AAA36487.1; -.
DR PIR: A48118; A48118.
DR HSSP: P02593; 1CDM.
DR GO: GO:0005856; C:cytoskeleton; NAS.
DR GO: GO:0005509; F:calcium ion binding; NAS.
DR GO: GO:0003154; P:cell differentiation; NAS.
DR GO: GO:0008151; P:cell growth and/or maintenance; NAS.
DR InterPro: IPR001751; CAP_S100.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF03516; Filaggrin; 6.
DR Pfam: PF01023; S_100; 1.
DR PRINTS: PR00487; FILAGGRIN.
DR PROSITE: PS00018; EF_HAND. 1.
DR PROSITE: PS00303; S100_CAP; 1.
KW Phosphorylation; Polypeptide; Developmental protein; Calcium-binding;
KW Polymorphism.
FT CA BIND 19 32 SITE I (BY SIMILARITY).
FT CA BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 1218 1218
SQ SEQUENCE 1218 AA; 13604 MW; EC195AD5285B19C2 CRC64;

Query Match 74.8%; Score 86; DB 4; Length 1218;
Best Local Similarity 81.0%; Pred. No. 5,6e-05;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 HOCHBESTXGRSGRCGRSGS 21
DB 447 HOCHBESTXGRSGRCGRSGS 467

RESULT 4
Q9H4U3 ID 09H4U3 PRELIMINARY; PRT; 798 AA.
AC 09H4U3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE D14N1.1.2 (Profilaggrin 3' end) (Fragment).
GN FIG.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL356504; CAC13171.1; -.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR003303; Filaggrin.
DR Pfam: PF03516; Filaggrin; 4.
DR PRINTS: PR00487; FILAGGRIN.
FT NON TER 1
SQ SEQUENCE 798 AA; 84773 MW; F923DDA8D1290805 CRC64;

Query Match 68.7%; Score 79; DB 4; Length 798;
Best Local Similarity 80.0%; Pred. No. 0.00047;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 QCHBESTXGRSGRCGRSGS 21
DB 427 QCHBESTXGRSGRCGRSGS 446

RESULT 5
Q01212 ID 001212 PRELIMINARY; PRT; 1084 AA.

```

```

AC Q01212, Q03840;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Profilaggrin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91064447; PubMed=2248957;
RA Gan S.O., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published extratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60501; AAA63243.1; -.
DR EMBL; M60501; AAA63243.1; JOINED.
DR GO; GO:0005882; C:intermediate filament; NNS.
DR GO; GO:0005198; F:structural molecule activity; NNS.
DR GO; GO:0007275; P:development; NNS.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF03516; Filaggrin; 6.
DR PRINTS; PR00487; Filaggrin.
FT NON TER
SQ SEQUENCE 1084 AA; 115271 MW; 80C4640B8D5A362D CRC64;

Query Match 68.7%; Score 79; DB 4; Length 1084;
Best Local Similarity 80.0%; Pred. No. 0.0063;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 QCHOESTYGRSGRCGRSGS 21
Db 63 QSHOESTGRSGRCGRSGS 82

RESULT 6
Q01720 PRELIMINARY; PRT; 591 AA.
ID Q01720;
AC Q01720;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FILAGGRIN precursor (PROFILAGGRIN) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93054736; PubMed=1429717;
RA Praeland R.B., Haydock P.V., Fleckman P., Nurusukeiri W., Dale B.A.;
RT "Characterization of the human epidermal profilaggrin gene. Genomic
RT organization and identification of an S-100-like calcium binding
RT domain at the amino terminus."
RT J. Biol. Chem. 267:23772-23781(1992).
[1]
-1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE FILAMENTS DURING
TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
-1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF
324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES.
DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
PROBOLYTAILALLY CLEAVED
-1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FILAGGRIN
REPEATS.
EMBL; L01089; AAA60177.1; -.
EMBL; L01090; AAA60176.1; -.

```

```

DR PIR; A45135; A45135.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR MIM; 135940; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CapB_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 2.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CAPB; 1.
KW Polymorphism.
FT PROPEP 1 293
FT CHAIN 294 467 POTENTIAL.
FT PROPEP 468 474 POTENTIAL.
FT CHAIN 475 >591 FILAGGRIN.
FT CA_BIND 19 32 SITE I (BY SIMILARITY).
FT CA_BIND 62 73 SITE II (BY SIMILARITY).
FT NON TER 591 591
SQ SEQUENCE 591 AA; 66366 MW; 381491625C75E369 CRC64;

Query Match 60.9%; Score 70; DB 4; Length 591;
Best Local Similarity 75.0%; Pred. No. 0.0092;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 QCHOESTYGRSGRCGRSGS 21
Db 448 QSHOESTGRSGRCGRSGS 467

RESULT 7
Q9H402 PRELIMINARY; PRT; 687 AA.
ID Q9H402;
AC Q9H402;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE D014N1.1.1 (Profilaggrin 5' end) (Fragment).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AL356504; CAC13172.1; -.
DR PIR; A48118; A48118.
DR HSSP; P80511; 1E8A.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001751; CapB_S100.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003303; Filaggrin.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF03516; Filaggrin; 3.
DR Pfam; PF01023; S_100; 1.
DR PRINTS; PR00487; FILAGGRIN.
DR PROSITE; PS00018; EF_HAND_1.
DR PROSITE; PS00303; S100_CAPB; 1.
FT NON TER 687 687
SQ SEQUENCE 687 AA; 76659 MW; 8000363FBEF07B74 CRC64;

Query Match 60.9%; Score 70; DB 4; Length 687;
Best Local Similarity 75.0%; Pred. No. 0.011;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      2 QCHOESTXGRSGRSGSGS 21
      |||||
DB      448 QSHOESTRGRSGRSGSGS 467

RESULT 8
ID      003838      PRELIMINARY;      PRT;      465 AA.
AC      Q03838;
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT      01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE      FLAGGRIN (PROFLAGGRIN) (Fragment).
OS      FIG.
GN      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC      NCBI_TaxID=9606;
RP      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RX      MEDLINE=91064347; PubMed=2248957;
RX      Gen S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
RT      gene."
RL      Biochemistry 23:9432-9440 (1990).
RN      [2]
RP      REVISIONS.
RX      MEDLINE=91255199; PubMed=2043621;
RX      Gen S.-Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT      "Organization, structure, and polymorphisms of the human profilaggrin
RT      gene."
RL      Biochemistry 30:5814-5814 (1991).
CC      -1- FUNCTION: FLAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
CC      PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC      FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC      -1- POLYMORPHISM: A NUMBER OF PROFILAGGRIN ISOFORMS HAVE BEEN FOUND
CC      WHICH DIFFER BOTH IN SEQUENCE AND IN THE NUMBER OF FLAGGRIN
CC      REPEATS.
CC      -1- MISCELLANEOUS: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE,
CC      INSOLUBLE, HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM
CC      COPIES OF 324 AA. THE PRECURSOR IS DEPOSITED AS KERATOHyalin
CC      GRANULES. DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC      PROTEOLYTICALLY CLEAVED.
DR      EMBL; M60499; AAA63246.1; -
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR003303; Flaggrin.
DR      Pfam; PF03516; Flaggrin; 3.
DR      PRINTS; PR00487; FLAGGRIN.
FT      NON_TER      1
FT      NON_TER      465
SQ      SEQUENCE 465 AA; 50280 MW; C883744C5E134097 CRC64;

Query Match      56.5%; Score 65; DB 4; Length 465;
Best Local Similarity 70.0%; Pred. No. 0.045;
Matches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 QCHOESTXGRSGRSGSGS 21
      |||||
DB      226 QSHOESARGRSGRSGSGS 245

RESULT 9
ID      075370      PRELIMINARY;      PRT;      322 AA.
AC      075370;
DT      01-NOV-1998 (TREMblrel. 08, Created)
DT      01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT      01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE      Epidermal filaggrin (Fragment).
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=99101527; PubMed=9886436;
RX      Girbal-Neuhausser B., Durlieu J.J., Arnaud M., Dalbon P., Sebbag M.,
RA      Vincent C., Simon M., Sersu T., Maeson-Bessiere C.,
RA      Jolivet-Reynaud C., Jolivet M., Serre G.;
RT      "The epitopes targeted by the rheumatoid arthritis-associated
RT      anti-flaggrin autoantibodies are posttranslationally generated on
RT      various sites of (pro)flaggrin by delamination of arginine residues."
RL      J. Immunol. 162:585-594 (1999).
DR      EMBL; AF043380; AAC23559.1; -
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      InterPro; IPR003303; Flaggrin.
DR      Pfam; PF03516; Flaggrin; 2.
DR      PRINTS; PR00487; FLAGGRIN.
FT      NON_TER      1
FT      NON_TER      322
SQ      SEQUENCE 322 AA; 34084 MW; 0DC2D0230D8FF9E0 CRC64;

Query Match      55.7%; Score 64; DB 4; Length 322;
Best Local Similarity 68.4%; Pred. No. 0.045;
Matches 13; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      2 QCHOESTXGRSGRSGSGS 20
      |||||
DB      304 QSHOEFTRGSGRSGSGS 322

RESULT 10
ID      09CWS6      PRELIMINARY;      PRT;      305 AA.
AC      09CWS6;
DT      01-JUN-2001 (TREMblrel. 17, Created)
DT      01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT      01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE      ES cells cDNA, RIKEN full-length enriched library, clone:2410005016
DE      product:similar to CGI-90 protein.
GN      2410005016RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RA      Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA      Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA      Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA      Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA      Kawai U., Kojima Y., Komio H., Kouda M., Koya S., Kurihara C.,
RA      Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA      Okazaki Y., Okido T., Owa C., Saito R., Saito R., Sakai C., Sakai K.,
RA      Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA      Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA      Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA      Yamatsuta M., Hayashizaki Y.,
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium.
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs."
RL      Nature 420:563-573 (2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RX      MEDLINE=21085660; PubMed=11217851;
RA      RIKEN FANTOM Consortium;
RT      "Functional annotation of a full-length mouse cDNA collection."

```



```

RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama U., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multichipillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
DR EMBL; AK010421; BAB26924.2; -
DR MGD; MGI:1913552; 241005016Rik.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
SQ SEQUENCE 305 AA; 35085 MW; 9EFF3012775F766 CRC64;

Query Match 46.1%; Score 53.5; DB 11; Length 305;
Best Local Similarity 66.7%; Pred. No. 1.9;
Matches 12; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 3 CHOESTXGSRGRCGRSG 20
DB 17 COQRTWG-SRGRCGRWG 33

RESULT 11
OQC297 PRELIMINARY; PRT; 113 AA.
AC OQC297;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 2810032G03Rik protein.
GN 2810032G03Rik
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shingawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Stadhil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

RA Brownstein M.J., Bul C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni U., Mashima U., Mazzarelli U., Monbets P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK012851; BAB28512.1; -
DR MGD; MGI:1919919; 2810032G03Rik.
SQ SEQUENCE 113 AA; 13026 MW; A4DE33BEED8F143 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 113;
Best Local Similarity 52.4%; Pred. No. 0.9;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 HOCHOESTXGSRGRCGRSG 21
DB 3 HPCHESSVWRSMRRLGLGS 23

RESULT 12
OQB722 PRELIMINARY; PRT; 113 AA.
AC OQB722;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN 2810032G03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028030; BAC25709.1; -
DR MGD; MGI:1919919; 2810032G03Rik.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 13010 MW; A4DE33BEF38F143 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 113;
Best Local Similarity 52.4%; Pred. No. 0.9;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 HOCHOESTXGSRGRCGRSG 21
DB 3 HPCHESSVWRSMRRLGLGS 23

RESULT 13
OQBDR9 PRELIMINARY; PRT; 130 AA.
AC OQBDR9;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Hypothetical protein.
GN 2810032G03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/60; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK029678; BAC26560.1; -;
 DR MGD; MGI:1919919; 2810032G03Rik.
 KM Hypothetical protein.
 SQ SEQUENCE 130 AA; 14981 MW; AE5669C54B09D234 CRC64;

Query Match 46.1%; Score 53; DB 11; Length 130;
 Best Local Similarity 52.4%; Pred. No. 1;
 Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

DQ 1 HCHQESTXGRSGRCRSGS 21
 20 HPCHESSVWRSMRRLGLGS 40

RESULT 14
 Q9P312 PRELIMINARY; PRT; 822 AA.
 AC Q9P312;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Related to nucleolar phosphoprotein.
 GN B12FL10.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algn V., Hohnselt J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL390091; CAB98213.1; -;
 DR PIR; TS1049; TS1049.
 DR GO; GO:0003676; Functional acid binding; IEA.
 DR InterPro; IPR000504; RNA rec. mot.
 DR InterPro; IPR005120; Smg-4_DPF3.
 DR Pfam; PF00076; Rrm; 1.
 DR Pfam; PF03467; Smg4_DPF3; 1.
 DR SMART; SMO0360; RRM; 1.
 DR PROSITE; PSS0102; RRM; 1.
 DR PROSITE; PSS0030; RRM_RNP_1; 1.
 SQ SEQUENCE 822 AA; 86287 MW; E40A457DC077245C CRC64;

Query Match 46.1%; Score 53; DB 3; Length 822;
 Best Local Similarity 58.8%; Pred. No. 5.9;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DQ 5 QESTXGRSGRCRSGS 21
 414 RESASGTRGRGRGGT 430

RESULT 15
 Q8B489 PRELIMINARY; PRT; 947 AA.
 AC Q8B489;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NS5 protein (Fragment).

GN NS5.
 OS Dengue virus type 2.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=11060;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Baleotti F.G., Koreli M.L., Figueiredo L.T.M.;
 RT "Detection of RNA polymerase motifs in the NS5 protein of Brazilian
 RT Flavivirus.";
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY167440; AAN7510.1; -;
 DR GO; GO:0005524; F-ATP binding; IEA.
 DR GO; GO:0003958; F-RNA-directed RNA polymerase activity; IEA.
 DR InterPro; IPR00208; Flavi_NS5.
 DR InterPro; IPR002877; RrmU_FtsJ.
 DR Pfam; PF00972; Flavi_NS5; 1.
 DR Pfam; PF01728; FtsJ; 1.
 DR NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 947 AA; 109981 MW; A644561B2864F8DC CRC64;

Query Match 46.1%; Score 53; DB 12; Length 947;
 Best Local Similarity 56.2%; Pred. No. 6.8;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

DQ 1 HCHQESTXGRSGRC 16
 849 HAHEH1QKGRGRGC 864

Search completed: September 28, 2004, 06:12:51
 Job time : 47.4052 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 59.8281 Seconds

Title: US-09-308-150-10

Sequence: 1 HQCHQESTXGRSRGRCGRSGS 21

Scoring table: BIOSUM62

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```
Database :
1: Genesegp.29Jan04.*
2: Genesegp.1980s.*
3: Genesegp.1990s.*
4: Genesegp.2000s.*
5: Genesegp.2001s.*
6: Genesegp.2002s.*
7: Genesegp.2003s.*
8: Genesegp.2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	113	98.3	21	2	AAW61520	AAW61520 Peptide X
2	113	98.3	22	4	AAE07235	AAE07235 IGF1546 P
3	86	74.8	1467	5	ABB97605	ABB97605 Novel hum
4	82	71.3	19	2	AAW61515	AAW61515 Peptide c
5	81	70.4	19	2	AAW61517	AAW61517 Peptide c
6	81	70.4	19	2	AAW61505	AAW61505 Peptide c
7	81	70.4	19	2	AAW61514	AAW61514 Peptide c
8	81	70.4	19	2	AAW61516	AAW61516 Peptide c
9	75	65.2	19	2	AAW61506	AAW61506 Peptide c
10	75	65.2	19	2	AAW61508	AAW61508 Peptide c
11	75	65.2	19	2	AAW61507	AAW61507 Peptide c
12	75	65.2	19	2	AAW61512	AAW61512 Peptide c
13	75	65.2	19	2	AAW61509	AAW61509 Peptide c
14	75	65.2	19	2	AAW61511	AAW61511 Peptide c
15	75	65.2	19	2	AAW61513	AAW61513 Peptide c
16	75	65.2	19	2	AAW61510	AAW61510 Peptide c
17	64	55.7	330	2	AAW22954	AAW22954 Human fl1
18	54	47.4	201	4	AAW79771	AAW79771 Human pro
19	54	47.0	70	5	ABP00440	ABP00440 Human ORF
20	54	47.0	330	2	AAW22956	AAW22956 Human fl1
21	54	47.0	330	2	AAW22955	AAW22955 Human fl1
22	54	47.0	330	2	AAW22957	AAW22957 Human fl1
23	52	45.2	424	4	ABW70938	ABW70938 Drosophila
24	50	43.5	477	6	AAW07142	AAW07142 Novel hum
25	50	43.5	1711	4	AAW79819	AAW79819 Human pro

27	50	43.5	1.551	4	AAAT78835	Aam78835 Human prot
26	48.5	42.2	4.88	2	AAAT77757	Aa67757 Human fet
25	48.5	42.2	5.24	2	AAAT67758	Aa67758 Human fet
28	48.5	42.2	6.70	2	AAAT67759	Aa67759 Human fet
29	48.5	42.2	7.69	2	AAAT53552	Aa75352 Human fet
30	48.5	42.2	8.57	2	AAAT53557	Aa75357 Human fet
31	48.5	42.2	8.61	7	ABD98487	Abd98487 Novel prot
32	48.5	41.7	21.4	5	ABAT92882	Abb92882 Herbi-cida
33	48.5	41.7	6.41	4	ABSL13345	Abg13945 Novel hum
33	47.5	41.3	6.54	7	ADDA46732	Kat Prote
33	47.5	41.3	8.54	7	ADDA46736	Kat Prote
35	47.5	41.3	8.54	7	ADDA46736	Rat Prote
36	47.5	41.3	8.54	7	ADDA55432	Rat Prote
37	47.5	41.3	8.54	7	ADDA55432	Rat Prote
38	46.5	40.9	12.8	5	ABBO47116	Abbo47116 Human p
38	46.5	40.9	12.8	5	ABBO47116	Human p
37	46.5	40.4	13.6	4	AAU282865	Aau28286 Novel hum
37	46.5	40.4	13.6	4	AAU282865	Novel hum
39	46.5	40.0	10.5	6	AAU61611	Propionib
39	46.5	40.0	10.5	6	ABMS59130	Propionib
41	45	39.1	5.2	4	AAU56816	Propionib
41	45	39.1	5.2	6	ABMS53335	Propionib
42	45	39.1	1.86	7	ADBA64751	Human pro
43	45	39.1	2.11	5	ABU96183	Human sec
44	45	39.1	2.11	5	ABU96183	Human sec
45	45	39.1	2.11	5	ABG64841	Human alb

ALIGNMENTS

RESULT 1

ID AAW61520 standard; peptide; 21 AA.

DT 26-OCT-1998 (first entry)

Peptide XI based on cDNA of a profilaggrin repeat

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody; rheumatoid arthritis; polyarthritis; antibody.

monoclonal antibody.

Synthetic.

	Kor	XX	XX	Location/Qualifiers
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

FT	Modified-site	9	/note-	"Cittm11ine"
FT				

XX
PN
W09822503-A2

XX 28-MAY-1998.
PD

XX 14-NOV-1997: 97WO-NL000624.
PF

XX 15-NOV-1996: 96NT-01004539.
PR

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER
PA

FA (LEWE-) SICHTUNG ICH WEIENSCHAFEN.
XX

XX
F1 VAN VENHOOF, HON, SCHEFFERS ON, ADAMS ON, 1900 1900

NET 1 220-2200/23
XX

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

[illegible]

XX

CC the profilaggrin antigen which is recognised by autoantibodies from

RA patient's autoimmune antibodies which are reactive with profilaggrin.

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 XX
 SQ Sequence 21 AA;

Query Match 98.3%; Score 113; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 1 HCHQESTXGRSGRCGRSGS 21

RESULT 2
 AAB07235
 ID AAB07235 standard; peptide; 22 AA.
 XX
 AC AAB07235;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE IGP1546 peptide for diagnosis and treatment of rheumatoid arthritis.
 XX
 KM Synthetic peptide; IGP1546; autoimmune antibody; rheumatoid arthritis;
 KM therapy; autoimmune disease; antirheumatic; systemic hyporesponsiveness;
 KM immunosuppressive; antiarthritic; cyclic.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 4..17
 FT Modified-site 10
 FT /note= "Citrulline"
 XX
 PN WO200146222-A2.
 XX
 PD 28-JUN-2001.
 XX
 PF 20-DEC-2000; 2000WO-EP013037.
 XX
 PR 21-DEC-1999; 99EP-00870280.
 PR 08-SEP-2000; 2000EP-00870195.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Union A, Moereels H, Meheus L;
 XX
 DR WPI; 2001-49657/54.
 XX
 PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
 PT comprises citrulline residue between 2 cysteine residues and is
 PT specifically recognized by autoimmune antibodies from patients suffering
 PT from rheumatoid arthritis.
 XX
 PS Example 3; Page 32; 53pp; English.
 XX
 CC The present sequence is a synthetic peptide, IGP1546. The peptide of the
 CC invention comprises a citrulline residue between 2 cysteine residues and
 CC is specifically recognised by autoimmune antibodies from patients
 CC involved in rheumatoid arthritis. The peptide comprises amino acids
 CC involved in side chain interactions which is essential for the formation
 CC of three-dimensional structure of the peptide. The peptide of the
 CC invention is useful as a medicament to treat autoimmune diseases,
 CC preferably rheumatoid arthritis. It is also useful for treating
 CC autoimmune diseases by increasing the size of antigen-immune complexes to
 CC improve clearance of the formed immune complexes and for the preparation
 CC of a medicament for oral or nasal administration to treat autoimmune
 CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
 CC to the peptide
 XX
 SQ Sequence 22 AA;

Query Match 99.3%; Score 113; DB 4; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 2 HCHQESTXGRSGRCGRSGS 22

RESULT 3
 AAB97605
 ID AAB97605 standard; protein; 1467 AA.
 XX
 AC AAB97605;
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE Novel human protein SEQ ID NO: 873.
 XX
 KM Human; antianaemic; vulnerary; antiinflammatory; immunomodulatory;
 KM antiinfectility; cerebroprotective; cytosratic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 KM expressed Sequence tag.
 XX
 OS Homo sapiens.
 XX
 PN WO200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-US026015.
 XX
 PR 11-SEP-2000; 2000US-00659671.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Aseudi V, Zhang J, Zhao QA, Ren F;
 PI Xue J, Yang Y, Wehrman T, Dimanac RT;
 XX
 DR WPI; 2002-292408/33.
 DR N-PSDB; ABN32791.
 XX
 PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 XX
 PS Example 2; SEQ ID NO 873; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 44
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX
 SQ Sequence 1467 AA;

Query Match 74.8%; Score 86; DB 5; Length 1467;
 Best Local Similarity 81.0%; Pred. No. 0.00059;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 HCHQESTXGRSGRCGRSGS 21
 |||||
 DB 447 HCHQESTXGRSGRCGRSGS 467

RESULT 4
 AAM61515
 ID AAM61515 standard; peptide; 19 AA.
 XX

```

AC AAW61515;
XX
XX 26-OCT-1998 (first entry)
DT
XX
XX Peptide cfa, based on cDNA of a proflilagrin repeat.
DE
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX solid phase synthesis; peptide amide; polyclonal antibody;
XX monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the proflilagrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
Query Match 71.3%; Score 82; DB 2; Length 19;
Best Local Similarity 88.9%; Pred. No. 3.4e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 HOESTXGRSRGRCGRSGS 21
Db ||||| ||||| |||||
2 HOESTXGRSRGRCGRSGS 19
RESULT 5
AAW61517
ID AAW61517 standard; peptide; 19 AA.
XX
XX AAW61517;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfo, based on cDNA of a proflilagrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KW solid phase synthesis; peptide amide; polyclonal antibody;
XX KW monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09822503-A2.
XX

```

```

PD 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX WPI; 1998-398613/34.
XX
XX Peptide derived from an antigen recognised by autoantibodies - is
XX PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
XX PT used in diagnosis of the disease.
XX
XX Disclosure; Page 6; 19pp; English.
XX
XX Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
XX CC the proflilagrin antigen which is recognised by autoantibodies from
XX CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
XX CC RA patient's autoimmune antibodies which are reactive with proflilagrin.
XX CC The peptides were created by using standard solid phase synthesis, which
XX CC produced them as peptide amides. These sequences may be used in the
XX CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
XX CC for obtaining polyclonal and monoclonal antibodies
XX
XX Sequence 19 AA;
SQ
Query Match 70.4%; Score 81; DB 2; Length 19;
Best Local Similarity 88.9%; Pred. No. 4.7e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 HOESTXGRSRGRCGRSGS 21
Db ||||| ||||| |||||
2 HOESTXGRSRGRCGRSGS 19
RESULT 6
AAW61505
ID AAW61505 standard; peptide; 19 AA.
XX
XX AAW61505;
XX
XX 26-OCT-1998 (first entry)
XX
XX Peptide cfo1, based on cDNA of a proflilagrin repeat.
XX
XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
XX KW solid phase synthesis; peptide amide; polyclonal antibody;
XX KW monoclonal antibody.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Modified-site 7 /note="Citruiline"
XX
XX W09822503-A2.
XX
XX 28-MAY-1998.
XX
XX 14-NOV-1997; 97WO-NL000624.
XX
XX 15-NOV-1996; 96NL-01004539.
XX
XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;
XX

```

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 94.4%; Pred. No. 4.7e-05;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTXGRSGRCGRSGS 19

RESULT 7

AAW61514

ID AAW61514 standard; peptide; 19 AA.

XX

AC AAW61514;

XX

XX 26-OCT-1998 (first entry)

XX

XX Peptide cf, based on cDNA of a proflilagrin repeat.

DE

XX

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KM solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX W09822503-A2.

PN

XX 28-MAY-1998.

PD

XX

XX 14-NOV-1997; 97WO-NL000624.

PF

XX

XX 15-NOV-1996; 96NL-01004539.

PR

XX

PA (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

PA (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

PI WPI; 1998-398613/34.

DR

XX

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 88.9%; Pred. No. 4.7e-05;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTXGRSGRCGRSGS 19

RESULT 8

AAW61516

ID AAW61516 standard; peptide; 19 AA.

XX

AC AAW61516;

XX

XX 26-OCT-1998 (first entry)

XX

XX Peptide cfE, based on cDNA of a proflilagrin repeat.

DE

XX

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KM solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclonal antibody.

XX

OS Synthetic.

OS Homo sapiens.

XX

XX W09822503-A2.

PN

XX 28-MAY-1998.

PD

XX

XX 14-NOV-1997; 97WO-NL000624.

PF

XX

XX 15-NOV-1996; 96NL-01004539.

PR

XX

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX

XX Van Venrooij WJW, Schellekens GA, Raats JMH, Hoet RMA;

PI WPI; 1998-398613/34.

DR

XX

XX Peptide derived from an antigen recognised by autoantibodies - is

PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be

PT used in diagnosis of the disease.

XX

PS Disclosure; Page 6; 19pp; English.

XX

CC Sequences AAW61505-W61520 are peptides derived from the C-terminal end of

CC the proflilagrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a

CC RA patient's autoimmune antibodies which are reactive with proflilagrin.

CC The peptides were created by using standard solid phase synthesis, which

CC produced them as peptide amides. These sequences may be used in the

CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as

CC for obtaining polyclonal and monoclonal antibodies

XX

XX Sequence 19 AA;

XX

Query Match 70.4%; Score 81; DB 2; Length 19;

Best Local Similarity 88.9%; Pred. No. 4.7e-05;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRCGRSGS 21

Db 2 HOESTXGRSGRCGRSGS 19

RESULT 9

AAW61506

XX	AAW61506 standard; peptide; 19 AA.
AC	AAW61506;
DT	26-OCT-1998 (first entry)
DE	Peptide cfc2, based on cDNA of a profilaggrin repeat.
XX	
KW	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody; solid phase synthesis; peptide amide; polyclonal antibody;
KX	monoclonal antibody.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 9
FT	/note="Citnulline"
XX	
PN	WO9822503-A2.
XX	
PD	28-MAY-1998.
XX	
PF	14-NOV-1997; 97MO-NL000624.
XX	
PR	15-NOV-1996; 96NL-01004539.
XX	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER. (TEWE-) STICHTING TECH WETENSCHAPPEN.
PI	Van Venrooij WJM, Schellekens GA, Raats JWH, Hoet RMA, WPI; 1998-398623/34.
DR	
XX	
PT	Peptide derived from an antigen recognised by autoantibodies - is reactive with autoimmune antibodies from rheumatoid arthritis, and may be used in diagnosis of the disease.
PT	
XX	
PS	Disclosure; Page 6; 19pp; English.
XX	
CC	Sequences AAW61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from RA patients with rheumatoid arthritis (RA). This peptide is reactive with a RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which produced them as peptide amides. These sequences may be used in the detection of RA autoimmune antibodies, in the diagnosis of RA, as well as for obtaining polyclonal and monoclonal antibodies
CC	
CC	
SO	Sequence 19 AA;
QY	Query Match 65.2%; Score 75; DB 2; Length 19; Best Local Similarity 83.3%; Pred. No. 0.00036; Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0
Dy	4 HOESTXGRSGRCGRSGS 21 2 HQSTRXSRGRCGRSGS 19
RESULT 10	
ID	AAW61508
ID	AAW61508 standard; peptide; 19 AA.
AC	AAW61508;
XX	
DT	26-OCT-1998 (first entry)
XX	
DE	Peptide cfc4, based on cDNA of a profilaggrin repeat.
XX	
KM	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody; solid phase synthesis; peptide amide; polyclonal antibody;
KM	monoclonal antibody.

XX	Synthetic.
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site
FT	Location/Qualifiers 13 /note= "Citru[ulline]"
XX	
FN	MO9822503-A2.
XX	
PD	28-MAY-1998.
XX	
PF	14-NOV-1997; 97WO-NL000624.
XX	
PR	15-NOV-1996; 96NL-01004539.
XX	
PA	(SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER. (TME-) STICHTING TECH WETENSCHAPPEN.
XX	
P1	Van Venrooij WJM, Schellekens GA, Raats JMH, Hoet RMA; WPI; 1998-398613/34.
DR	
PT	Peptide derived from an antigen recognised by autoantibodies - is reactive with autoimmune antibodies from rheumatoid arthritis, and may be used in diagnosis of the disease.
XX	
PS	Disclosure; Page 6; 19pp; English.
CC	
CC	Sequences AAM61505-W61520 are peptides derived from the C-terminal end of the profilaggrin antigen which is recognised by autoantibodies from patients with rheumatoid arthritis (RA). This peptide is reactive with a RA patient's autoimmune antibodies which are reactive with profilaggrin. The peptides were created by using standard solid phase synthesis, which produced them as peptide amides. These sequences may be used in the detection of RA autoimmune antibodies, in the diagnosis of RA, as well as for obtaining polyclonal and monoclonal antibodies
CC	
XX	
SQ	Sequence 19 AA;
Query Match	65.2%; Score 75; DB 2; Length 19;
Best Local Similarity	83.3%; Pred. No. 0.00036;
Matches 15; Conservative	0; Mismatches 3; Indels 0; Gaps 0
OY	4 HOESTXGRSRGRCGRSGS 21 Db 2 HOESTRGRSRGXSGRSGS 19
RESULT 11	
AAM61507	
ID	AAM61507 standard; peptide; 19 AA.
XX	
AC	AAM61507;
XX	
DT	26-OCT-1998 (first entry)
XX	
DE	Peptide cfc3, based on cDNA of a profilaggrin repeat.
XX	
KM	Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody; solid phase synthesis; peptide amide; polyclonal antibody; monoclonal antibody.
XX	
OS	Synthetic.
XX	
OS	Homo sapiens.
FH	Key
FT	Modified-site
FT	Location/Qualifiers 11 /note= "Citru[ulline]"
PN	WO9822503-A2.
XX	
XX	
PD	28-MAY-1998.

XX 14-NOV-1997; 97WO-NL000624.
 PF 15-NOV-1996; 96NL-01004539.
 PR (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 DR WPI; 1998-398613/34.
 XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PR used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 83.3%; Pred. No. 0.00036;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRCGRSGS 21
 |||||
 DB 2 HOESTRGRSGRCGRSGS 19

RESULT 12
 AAW61512
 ID AAW61512 standard; peptide; 19 AA.
 XX
 AC AAW61512;

DT 26-OCT-1998 (first entry)

DE Peptide cfc8, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FT Key Location/Qualifiers
 FT Modified-site 7 /note= "Citruilline"
 FT Modified-site 13 /note= "Citruilline"
 FT /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 PT

XX WPI; 1998-398613/34.
 DR

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PR used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from
 CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 CC
 CC Sequence 19 AA;
 SQ

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 88.9%; Pred. No. 0.00036;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSGRCGRSGS 21
 |||||
 DB 2 HOESTXGRSGRCGRSGS 19

RESULT 13
 AAW61509
 ID AAW61509 standard; peptide; 19 AA.
 XX
 AC AAW61509;

DT 26-OCT-1998 (first entry)

DE Peptide cfc5, based on cDNA of a profilaggrin repeat.

XX Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;
 KM solid phase synthesis; peptide amide; polyclonal antibody;
 KW monoclonal antibody.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

FT Key Location/Qualifiers
 FT Modified-site 16 /note= "Citruilline"
 FT /note= "Citruilline"

PN WO9822503-A2.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-NL000624.

PR 15-NOV-1996; 96NL-01004539.

XX (SCHE-) STICHTING SCHEIKUNDIG ONDERZOEK IN NEDER.
 PA (TEWE-) STICHTING TECH WETENSCHAPPEN.
 XX Van Venrooij MJW, Schellekens GA, Raats JMH, Hoet RMA;
 PT

DR WPI; 1998-398613/34.

XX Peptide derived from an antigen recognised by autoantibodies - is
 PT reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 PR used in diagnosis of the disease.
 XX Disclosure; Page 6; 19pp; English.
 PS Sequences AAW61505-W61520 are peptides derived from the C-terminal end of
 CC the profilaggrin antigen which is recognised by autoantibodies from

CC patients with rheumatoid arthritis (RA). This peptide is reactive with a
 CC RA patient's autoimmune antibodies which are reactive with profilaggrin.
 CC The peptides were created by using standard solid phase synthesis, which
 CC produced them as peptide amides. These sequences may be used in the
 CC detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 CC for obtaining polyclonal and monoclonal antibodies
 XX

Sequence 19 AA:

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 83.3%; Pred. No. 0.00036;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRSGSGS 21
 |||||
 DB 2 HOESTXGRSRGRSGSGS 19

RESULT 14

AAW61511
 ID AAW61511 standard; peptide; 19 AA.

XX AAW61511;

DT 26-OCT-1998 (first entry)

DE Peptide cfc7, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

OS Synthetic.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequence AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.
 The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the
 detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Query Match 65.2%; Score 75; DB 2; Length 19;
 Best Local Similarity 88.3%; Pred. No. 0.00036;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 HOESTXGRSRGRSGSGS 21
 |||||
 DB 2 HOESTXGRSRGRSGSGS 19

RESULT 15

AAW61513
 ID AAW61513 standard; peptide; 19 AA.

XX AAW61513;

DT 26-OCT-1998 (first entry)

DE Peptide cfc9, based on cDNA of a profilaggrin repeat.

KW Antigen; autoantibody; rheumatoid arthritis; RA; autoimmune antibody;

KW solid phase synthesis; peptide amide; polyclonal antibody;

XX monoclinal antibody.

OS Synthetic.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Peptide derived from an antigen recognised by autoantibodies - is
 reactive with autoimmune antibodies from rheumatoid arthritis, and may be
 used in diagnosis of the disease.

Disclosure; Page 6; 19pp; English.

Sequence AAW61505-W61520 are peptides derived from the C-terminal end of
 the profilaggrin antigen which is recognised by autoantibodies from
 patients with rheumatoid arthritis (RA). This peptide is reactive with a
 RA patient's autoimmune antibodies which are reactive with profilaggrin.
 The peptides were created by using standard solid phase synthesis, which
 produced them as peptide amides. These sequences may be used in the
 detection of RA autoimmune antibodies, in the diagnosis of RA, as well as
 for obtaining polyclonal and monoclonal antibodies

Search completed: September 28, 2004, 06:24:22
Job time : 61.8281 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:15:16 / Search time 123.266 Seconds
(without alignments)
54.782 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HQCHQSTXGRSRGRCGRSGS 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	98.3	21	9	US-09-308-150-10
2	113	98.3	22	9	US-09-747-029A-22
3	82	71.3	19	9	US-09-308-150-12
4	81	70.4	19	9	US-09-308-150-11
5	81	70.4	19	9	US-09-308-150-11
6	81	70.4	19	9	US-09-308-150-13
7	81	70.4	19	9	US-09-308-150-14
8	75	65.2	19	9	US-09-308-150-2
9	75	65.2	19	9	US-09-308-150-3
10	75	65.2	19	9	US-09-308-150-4
11	75	65.2	19	9	US-09-308-150-5
12	75	65.2	19	9	US-09-308-150-6
13	75	65.2	19	9	US-09-308-150-7
14	75	65.2	19	9	US-09-308-150-8
15	75	65.2	19	9	US-09-308-150-9

16	52	45.2	172	16	US-10-437-963-175440	Sequence 175440,
17	50	43.5	206	16	US-10-437-963-110381	Sequence 110381,
18	49	43.5	477	15	US-10-161-927-62	Sequence 62, Appl
19	49	42.6	103	16	US-10-437-963-122475	Sequence 122475,
20	49	42.6	110	16	US-10-437-963-104204	Sequence 104204,
21	49	42.6	150	16	US-10-437-963-172443	Sequence 172443,
22	48.5	42.2	88	16	US-10-437-963-131735	Sequence 131735,
23	48.5	42.2	160	12	US-10-425-114-66342	Sequence 66342, A
24	48.5	42.2	526	15	US-10-369-993-10496	Sequence 10496, A
25	48	41.7	1087	16	US-10-437-963-123162	Sequence 123162,
26	48	41.7	1087	9	US-09-918-909-24	Sequence 24, Appl
27	48	41.7	1087	16	US-10-641-991-24	Sequence 24, Appl
28	47.5	41.3	176	16	US-10-437-963-156023	Sequence 156023,
29	47	40.9	201	16	US-10-437-963-174936	Sequence 174936,
30	47	40.9	218	12	US-10-425-114-68534	Sequence 48534, A
31	47	40.9	363	16	US-10-437-963-151018	Sequence 151018,
32	47	40.9	384	12	US-10-425-114-72136	Sequence 72136, A
33	47	40.9	384	12	US-10-425-114-72137	Sequence 72137, A
34	46.5	40.4	56	16	US-10-437-963-155194	Sequence 155194,
35	46.5	40.4	136	12	US-10-221-278-622	Sequence 622, App
36	46.5	40.4	136	15	US-10-231-172-622	Sequence 622, App
37	46.5	40.4	456	16	US-10-437-963-155551	Sequence 155551,
38	46	40.4	53	16	US-10-767-701-57837	Sequence 57837, A
39	46	40.0	108	16	US-10-437-963-147392	Sequence 147392,
40	46	40.0	506	16	US-10-437-963-185974	Sequence 185974,
41	46	40.0	807	12	US-10-425-114-43305	Sequence 43305, A
42	46	40.0	807	12	US-10-425-114-45403	Sequence 45403, A
43	46	40.0	840	16	US-10-437-963-151528	Sequence 151528,
44	45	39.1	42	14	US-10-029-386-31444	Sequence 31444, A
45	45	39.1	164	16	US-10-437-963-112419	Sequence 112419,

ALIGNMENTS

RESULT 1

US-09-308-150-10

Sequence 10, Application US/09308150

Patent No. US020137092A1

GENERAL INFORMATION:

APPLICANT: Van Vencroij, Waltherus Jacobus Wilhelmus

APPLICANT: Schellekens, Gerardus Antonius

APPLICANT: Raats, Jozef Maria Hendrik

APPLICANT: Hoeft, Rene Michael Antonius

APPLICANT: Stichting Scheikundig Onderzoek Nederland

APPLICANT: Stichting voor de Technische Wetenschappen

TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY

TITLE OF INVENTION: ANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,

TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL

FILE REFERENCE: 09/308,150 -- PCT/NL97/00624

CURRENT APPLICATION NUMBER: US/09/308,150

CURRENT FILING DATE: 1999-09-30

PRIOR APPLICATION NUMBER: PCT/NL97/00624

PRIOR FILING DATE: 1997-11-14

PRIOR APPLICATION NUMBER: NL 1004539

PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 21

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:Derived from

OTHER INFORMATION: known cDNA sequences of human profilaggrin

NAME/KEY: DISULFID

LOCATION: (3)..(16)

US-09-308-150-10

Query Match

Best Local Similarity

98.3%; Score 113; DB 9; Length 21;

100.0%; Pred. No. 1.2e-08;

```

; LENGTH: 13
; TYPE: PRT
;

```

OTHER INFORMATION: Xaa is citrulline
US-09-308-150-1

7 TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
;
; TITLE OF INVENTION: ANTIGEN, AND A METHOD OF DETECTING AUTO-IM

```

; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human prolactin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-2

Query Match      65.2%; Score 75; DB 9; Length 19;

```

Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21
Db 2 HOESTRGRSGRSGRSGS 19

RESULT 9

US-09-308-150-3
Sequence 3, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-3

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21
Db 2 HOESTRGRSGRSGRSGS 19

RESULT 10

US-09-308-150-4
Sequence 4, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: US/09308,150
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15

NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-4

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21
Db 2 HOESTRGRSGRSGRSGS 19

RESULT 11

US-09-308-150-5
Sequence 5, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius
APPLICANT: Stichting Scheikundig Onderzoek Nederland
TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
FILE REFERENCE: 09/308,150 -- PCT/NL97/00624
CURRENT FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: PCT/NL97/00624
PRIOR FILING DATE: 1997-11-14
PRIOR APPLICATION NUMBER: NL 1004539
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Derived from
OTHER INFORMATION: known cDNA sequences of human profilaggrin
OTHER INFORMATION: Xaa is citrulline
US-09-308-150-5

Query Match 65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HOESTXGRSGRCGRSGS 21
Db 2 HOESTRGRSGRSGRSGS 19

RESULT 12

US-09-308-150-6
Sequence 6, Application US/09308150
Patent No. US20020137092A1

GENERAL INFORMATION:

APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
APPLICANT: Schellekens, Gerardus Antonius
APPLICANT: Raats, Jozef Maria Hendrik
APPLICANT: Hoeft, Rene Michael Antonius

```

; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308.150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-6

Query Match      65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
4 HOESTXGRSGRCGRSGS 21
|||||
2 HOESTXGRSGRCGRSGS 19

RESULT 13
US-09-308-150-7
; Sequence 7, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308.150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-7

Query Match      65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
4 HOESTXGRSGRCGRSGS 21
|||||
2 HOESTXGRSGRCGRSGS 19
```

```

; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308.150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-8

Query Match      65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
4 HOESTXGRSGRCGRSGS 21
|||||
2 HOESTXGRSGRCGRSGS 19

RESULT 14
US-09-308-150-8
; Sequence 8, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308.150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Derived from
; OTHER INFORMATION: known cDNA sequences of human profilaggrin
; OTHER INFORMATION: Xaa is citrulline
US-09-308-150-9

Query Match      65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. No. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
4 HOESTXGRSGRCGRSGS 21
|||||
2 HOESTXGRSGRCGRSGS 19

RESULT 15
US-09-308-150-9
; Sequence 9, Application US/09308150
; Patent No. US20020137092A1
; GENERAL INFORMATION:
; APPLICANT: Van Venrooij, Waltherus Jacobus Wilhelmus
; APPLICANT: Schellekens, Gerardus Antonius
; APPLICANT: Raats, Jozef Maria Hendrik
; APPLICANT: Hoeft, Rene Michael Antonius
; APPLICANT: Stichting Scheikundig Onderzoek Nederland
; TITLE OF INVENTION: PEPTIDE DERIVED FROM AN ANTIGEN RECOGNIZED BY
; TITLE OF INVENTION: AUTOANTIBODIES FROM PATIENTS WITH RHEUMATOID ARTHRITIS,
; TITLE OF INVENTION: ANTIBODY DIRECTED AGAINST SAID PEPTIDE, A COMBINATORIAL
; FILE REFERENCE: 09/308.150 -- PCT/NL97/00624
; CURRENT FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/308.150
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: NL 1004539
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 19
```

```
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Derived from
/ OTHER INFORMATION: known cDNA sequences of human profilaggrin
/ OTHER INFORMATION: Xaa is citrulline
us-09-308-150-9
```

```
Query Match          65.2%; Score 75; DB 9; Length 19;
Best Local Similarity 88.9%; Pred. NO. 0.0017;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      4 HOESTXGSRGRGCGSGS 21
          |||||
Db       2 HOESTXGSRGRGCGSGS 19
```

Search completed: September 28, 2004, 07:28:52
Job time : 123.266 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2004, 06:04:22 ; Search time 15.9688 Seconds

(without alignments)
67.892 Million cell updates/sec

Title: US-09-308-150-10
Perfect score: 115
Sequence: 1 HQHQUESTXGRSGRCRSGS 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.5	43.0	234	4	US-09-252-991A-17302
2	49.5	42.6	211	4	US-09-252-991A-24706
3	48.5	42.2	236	4	US-09-252-991A-25980
4	48.5	42.2	488	1	US-08-243-542-1
5	48.5	42.2	488	1	US-08-477-407-1
6	48.5	42.2	488	1	US-08-484-355-1
7	48.5	42.2	524	1	US-08-243-542-2
8	48.5	42.2	524	1	US-08-477-407-2
9	48.5	42.2	524	1	US-08-484-355-2
10	48.5	42.2	670	1	US-08-243-542-3
11	48.5	42.2	670	1	US-08-477-407-3
12	48.5	42.2	670	1	US-08-484-355-3
13	48.5	42.2	769	1	US-08-243-542-4
14	48.5	42.2	769	1	US-08-477-407-4
15	48.5	42.2	769	1	US-08-484-355-4
16	46.0	40.0	255	4	US-09-252-991A-19598
17	45.5	39.1	105	4	US-09-252-991A-12394
18	45.5	39.1	231	4	US-09-252-991A-17218
19	45.5	39.1	267	4	US-09-252-991A-18758
20	45.5	39.1	352	4	US-09-252-991A-18590
21	45.5	39.1	363	4	US-09-252-991A-25132
22	45.5	39.1	406	4	US-09-252-991A-19857
23	45.5	39.1	421	4	US-09-252-991A-23236
24	45.5	39.1	1151	4	US-09-252-991A-21328
25	44.5	38.7	83	4	US-09-716-129-151
26	44.5	38.7	696	3	US-09-351-414-2
27	44.5	38.7	834	4	US-09-252-991A-17616

28	44.5	38.7	923	4	US-09-252-991A-27148	Sequence 27148, A
29	44.5	38.7	957	4	US-09-252-991A-20408	Sequence 20408, A
30	43.5	37.8	62	4	US-09-252-991A-32266	Sequence 32266, A
31	43.5	37.8	212	4	US-09-252-991A-29831	Sequence 29831, A
32	43.5	37.8	292	4	US-09-199-637A-419	Sequence 419, App
33	43.5	37.4	178	4	US-09-252-991A-23975	Sequence 23975, A
34	43.5	37.4	203	4	US-09-252-991A-26436	Sequence 26436, A
35	43.5	37.4	225	4	US-09-252-991A-17754	Sequence 17754, A
36	43.5	37.4	865	4	US-09-252-991A-18683	Sequence 18683, A
37	43.5	37.4	1041	4	US-09-252-991A-25637	Sequence 25637, A
38	43.5	37.4	1253	1	US-08-452-966B-12	Sequence 12, Appl
39	42.5	37.0	591	4	US-09-252-991A-28760	Sequence 28760, A
40	42.5	36.5	175	4	US-09-252-991A-19711	Sequence 19711, A
41	42.5	36.5	404	4	US-09-489-039A-11606	Sequence 11606, A
42	42.5	36.5	405	4	US-09-252-991A-20011	Sequence 20011, A
43	42.5	36.5	411	4	US-09-252-991A-22445	Sequence 22445, A
44	42.5	36.5	483	4	US-09-252-991A-19015	Sequence 19015, A
45	42.5	36.5	518	4	US-09-252-991A-25967	Sequence 25967, A

ALIGNMENTS

```

RESULT 1
US-09-252-991A-17302
; Sequence 17302, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17302
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17302

Query Match      43.0%; Score 49.5; DB 4; Length 234;
Best Local Similarity 45.5%; Pred. No. 5.4;
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY      1 HQHQUESTXG-----RSRRCRG 17
Db      76 HECRDSTRGHGDSRPAFRCG 97

RESULT 2
US-09-252-991A-24706
; Sequence 24706, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24706
; LENGTH: 211
; TYPE: PRT

```

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24706

Query Match 42.6%; Score 49; DB 4; Length 211;
Best Local Similarity 56.2%; Pred. No. 5.8;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 QBSTXGRSGRCGRSG 20
DB 42 RSSAGRSRRRCARAG 57

RESULT 3
US-09-252-991A-25980
Sequence 25980, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25980
LENGTH: 236
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25980

Query Match 42.2%; Score 48.5; DB 4; Length 236;
Best Local Similarity 52.4%; Pred. No. 7.7;
Matches 11; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 4 HOESTXGRS-----RCGRCS 19
DB 78 HRRHRCGRSRDQGRRCGRCS 98

RESULT 4
US-08-243-542-1
Sequence 1, Application US/08243542
Patent No. 5552526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-1

Query Match 42.2%; Score 48.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHQESTXGRSGRCGRSGS 21
DB 461 CYEKLVNVEGTERSCGRKGS 480

RESULT 5
US-08-477-407-1
Sequence 1, Application US/08477407
Patent No. 5631351
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-477-407-1

Query Match 42.2%; Score 48.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO6-STXGRSGRCGRSGS 21
DB 461 CYEKLNVETGRSGCRKGS 480

RESULT 6
US-08-484-355-1
Sequence 1, Application US/08484355
Patent No. 5705341
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Wordperfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:

TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-484-355-1

Query Match 42.2%; Score 48.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO6-STXGRSGRCGRSGS 21
DB 461 CYEKLNVETGRSGCRKGS 480

RESULT 7
US-08-243-542-2
Sequence 2, Application US/08243542
Patent No. 5552526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Wordperfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furiya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 524 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSGRCGRSGS 21
Db 468 CYEXLNVEGTGRCGRSGS 487

RESULT 8
US-08-477-407-2

Sequence 2, Application US/08477407

Patent No. 5631351

GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.

STREET: 2026 Rambling Road

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49008-1699

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage

COMPUTER: IBM PC/XT/AT Compatible

OPERATING SYSTEM: MS-DOS 5.0

SOFTWARE: WordPerfect 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,407

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,542

FILING DATE: 13-MAY-1994

APPLICATION NUMBER: JP 5-136602

FILING DATE: 14 MAY 1993

APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993

APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328

FILING DATE: 12 APRIL 1994

APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994

ATTORNEY/AGENT INFORMATION:

NAME: Teriyence F. Chapman

REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Futuya Case 1313

TELECOMMUNICATION INFORMATION:

TELEPHONE: (616) 381-1156

TELEFAX: (616) 381-5465

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 524 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

LIBRARY: human fetal brain cDNA library

US-08-477-407-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSGRCGRSGS 21
Db 468 CYEXLNVEGTGRCGRSGS 487

RESULT 9

US-08-484-355-2

Sequence 2, Application US/08484355

Patent No. 5705341

GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.

STREET: 2026 Rambling Road

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49008-1699

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 MB storage

COMPUTER: IBM PC/XT/AT Compatible

OPERATING SYSTEM: MS-DOS 5.0

SOFTWARE: WordPerfect 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,355

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,542

FILING DATE: 13-MAY-1994

APPLICATION NUMBER: JP 5-136602

FILING DATE: 14 MAY 1993

APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993

APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328

FILING DATE: 12 APRIL 1994

APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994

ATTORNEY/AGENT INFORMATION:

NAME: Teriyence F. Chapman

REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Futuya Case 1313

TELECOMMUNICATION INFORMATION:

TELEPHONE: (616) 381-1156

TELEFAX: (616) 381-5465

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 524 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

LIBRARY: human fetal brain cDNA library

US-08-484-355-2

Query Match 42.2%; Score 48.5; DB 1; Length 524;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHO2-STXGRSGRCGRSGS 21
Db 468 CYEXLNVEGTGRCGRSGS 487

DB 468 CYEKLNVETGRSGCRKGS 487

RESULT 10

US-08-243-542-3

Sequence 3, Application US/08243542

Patent No. 5552526

GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.

STREET: 2026 Rambling Road

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49008-1699

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage

COMPUTER: IBM PC/XT/AT Compatible

OPERATING SYSTEM: MS-DOS 5.0

SOFTWARE: Wordperfect 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/243,542

FILING DATE: 22 APRIL 1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-136602

FILING DATE: 14 MAY 1993

APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993

APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328

FILING DATE: 12 APRIL 1994

APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994

ATTORNEY/AGENT INFORMATION:

NAME: Terryence F. Chapman

REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Furuya Case 1313

TELEPHONE: (616) 381-1156

TELEFAX: (616) 381-5465

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 670 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE: Homo sapiens

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

LIBRARY: human fetal brain cDNA library

US-08-243-542-3

Query Match 42.2%; Score 48.5; DB 1; Length 670;

Best Local Similarity 45.0%; Pred. No. 22;

Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHOE-STXGSRGRCGRSGS 21

DB 468 CYEKLNVETGRSGCRKGS 487

RESULT 11

US-08-477-407-3

Sequence 3, Application US/08477407

Patent No. 5631351

GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.

STREET: 2026 Rambling Road

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49008-1699

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage

COMPUTER: IBM PC/XT/AT Compatible

OPERATING SYSTEM: MS-DOS 5.0

SOFTWARE: Wordperfect 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,407

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/243,542

FILING DATE: 13-MAY-1994

APPLICATION NUMBER: JP 5-136602

FILING DATE: 14 MAY 1993

APPLICATION NUMBER: JP 5-257455

FILING DATE: 22 SEPTEMBER 1993

APPLICATION NUMBER: JP 6-49904

FILING DATE: 23 FEBRUARY 1994

APPLICATION NUMBER: JP 6-73328

FILING DATE: 12 APRIL 1994

APPLICATION NUMBER: JP 6-84470

FILING DATE: 22 APRIL 1994

ATTORNEY/AGENT INFORMATION:

NAME: Terryence F. Chapman

REGISTRATION NUMBER: 32 549

REFERENCE/DOCKET NUMBER: Furuya Case 1313

TELEPHONE: (616) 381-1156

TELEFAX: (616) 381-5465

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 670 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE: Homo sapiens

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

LIBRARY: human fetal brain cDNA library

US-08-477-407-3

Query Match 42.2%; Score 48.5; DB 1; Length 670;

Best Local Similarity 45.0%; Pred. No. 22;

Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 CHOE-STXGSRGRCGRSGS 21

DB 468 CYEKLNVETGRSGCRKGS 487

RESULT 12

US-08-484-355-3

Sequence 3, Application US/08484355

Patent No. 5705341

GENERAL INFORMATION:

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: EMI, MITSURU

TITLE OF INVENTION: MDC PROTEINS AND DNAs

TITLE OF INVENTION: ENCODING THE SAME

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 670 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-484-355-3
Query Match 42.2%; Score 48.5; DB 1; Length 670;
Best Local Similarity 45.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 3 CHOE-STXGRSRRCGRSGS 21
DB 468 CYEKLNVGTERGSCGRKGS 487
RESULT 13
US-08-243-542-4
Sequence 4, Application US/08243542
Patent No. 552526
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,542
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Futuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain cDNA library
US-08-243-542-4
Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
QY 3 CHOE-STXGRSRRCGRSGS 21
DB 567 CYEKLNVGTERGSCGRKGS 586
RESULT 14
US-08-477-407-4
Sequence 4, Application US/08477407
Patent No. 563151
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,407
FILING DATE: 07-JUN-1995

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993
APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain CDNA library
US-08-477-407-4

Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHOE-STXGRSRGRCGRSGS 21
Db 567 CYEKLNVGTERGSCGRKGS 586

RESULT 15
US-08-484-355-4
Sequence 4, Application US/08484355
Patent No. 5705341
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: EMI, MITSURU
TITLE OF INVENTION: MDC PROTEINS AND DNAs
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL & TANIS P.C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,355
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/243,542
FILING DATE: 13-MAY-1994
APPLICATION NUMBER: JP 5-136602
FILING DATE: 14 MAY 1993

APPLICATION NUMBER: JP 5-257455
FILING DATE: 22 SEPTEMBER 1993
APPLICATION NUMBER: JP 6-49904
FILING DATE: 23 FEBRUARY 1994
APPLICATION NUMBER: JP 6-73328
FILING DATE: 12 APRIL 1994
APPLICATION NUMBER: JP 6-84470
FILING DATE: 22 APRIL 1994
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32 549
REFERENCE/DOCKET NUMBER: Furuya Case 1313
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal brain CDNA library
US-08-484-355-4

Query Match 42.2%; Score 48.5; DB 1; Length 769;
Best Local Similarity 45.0%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 3 CHOE-STXGRSRGRCGRSGS 21
Db 567 CYEKLNVGTERGSCGRKGS 586

Search completed: September 28, 2004, 06:26:47
Job time : 16.9688 secs

Blank sheet